REMARKS/ARGUMENTS

Claims 1-3, 7, 8, 12, 13 and 47-50 are pending and were previously undergoing examination on the merits. Claims 1 and 2 and 12 are amended herein. There are no new claims. Claim 46 has been newly canceled without prejudice. Claims 9, 10, 24, 25, 27, 29, 30, 33, 37, 38, and 46 were previously canceled without prejudice. Claims 4 - 6, 11, 14-23, 26, 28, 31, 32, 34-36, and 39-43 stand withdrawn from consideration. After entry of the amendments, claims 1-3, 7, 8, 12, 13 and 47-50 will be pending.

Support for the Amendments

Claim 1 has been amended to recite in part "a polypeptide having a subsequence at least 95% identical." Support for such subsequence subject matter is found *inter alia* in the specification in the paragraph spanning pages 15 and 16 and the first paragraph on p.16.

Claim 1 has also been amended to set forth "an epitope presenting domain having an amino acid sequence of between 5 and 350 amino acids in length consisting essentially of one cysteine to cysteine disulfide bonded loop of a pathogen wherein the loop encodes an epitope of the pathogen and wherein the epitope is non-native to PE domain Ib." Support for such is found *inter alia* in the specification in the paragraph spanning pages 28 and 29 and in the previous version of the claim.

Claim 2 has been amended to delete subject matter. Support for such is found *inter alia* in the previous version of the claim.

In view of the above, the Applicants believe the amendments to the claims add no new matter and respectfully request their entry.

Status of the Claims

Applicants thank the Examiner for reconsidering and withdrawing the previous rejections under 35 U.S.C. §112 first and second paragraphs as well as the rejections under 35 U.S.C. §102(e).

The previous rejection of claims 1-3, 7-8, 12-13, and 46-50 under 35 U.S.C. §103 over Wels et al. (U.S. Patent No. 6,498,233) and Pastan et al. (U.S. Patent No. 6,074,644) was maintained.

Claims 2,7, and 8 stand objected to under 35 U.S.C.§1.75(c) for alleged improper form.

Claims 1, 46, and 47 stand rejected as allegedly failing to satisfy the written description requirement of 35 U.S.C.§112, first paragraph.

Applicants respectfully respond to each of the above:

Response to the Rejection under 35 U.S.C. §103

Applicants stand corrected as to nature of the nucleic acid binding domain of the protein component of the Wels et al. nucleic acid transfer system. As noted by the Examiner, Wels et al. disclose Gal4 can be used as the nucleic acid binding domain of their nucleic acid transfer system and Gal4. Applicants thank the Examiner for helping to clarify the subject matter disclosed by Wels et al.

In view of same, Applicants respectfully rebut the *prima facie* case on several grounds:

A. The proposed combination does not satisfy all the elements of amended claim 1.

Claim 1 has been amended to recite in part:

> "an epitope presenting domain having an amino acid sequence of between 5 and 350 amino acids in length and consisting essentially of one cysteine to cysteine bonded loop of a pathogen wherein the loop encodes an epitope of the pathogen and wherein the epitope is non-native to PE domain ".

Firstly, the proposed combination with $Gal4_{(1-149)}$ lacks a cysteine to cysteine bonded loop <u>of a pathogen.</u> Yeast $Gal4_{(1-149)}$ contains no such loop. The $Gal4_{(1-149)}$ cysteines are bonded through Zn(II) via a $Zn(II)_2Cys_6$ cluster as taught by Pan et al. as evidenced most conveniently in the Abstract and at Figure 7 of Pan et al. *PNAS* 87:2077-2081 (1990)(enclosed).

Figure 6 of Pan et al., PNAS 86:3145-49 (1989) sets forth the amino acid sequence of $Gal4_{(1-149)}$ and shows that $Gal4_{(1-149)}$ contains no other cysteine residues than those involved in the $Zn(II)_2Cys_6$ cluster. (Kindly note that Fig. 6 is cited only for the amino acid sequence set forth therein. The figure otherwise does not correctly show the secondary structure of 6 cysteines in a $Zn(II)_2Cys_6$ cluster - the later Pan et al. 1990 publication corrects the secondary structural arrangement and sets forth the $Zn(II)_2Cys_6$ cluster-.)

Secondly, the base claim has been amended to recite *consisting essentially of*. The combination proposed by the Examiner would necessarily include the nucleic acid binding portion of Gal4 in addition to any epitope portion of the Gal4 nucleic acid binding protein and would not *consist essentially of* an epitope presenting domain.

B. The fact that references can be combined or modified is not sufficient to establish *prima facie* obviousness.

Firstly, the need to be served by the proposed combination is not in evidence. The Examiner contends that the 6,074,644 patent discloses cysteine to cysteine disulfide bonds between two polypeptide chains can be used to stabilize them. Assuming for the sake of argument that such an effect would be even applicable to nucleic acid binding proteins generally, or Gal4 in particular, there is no indication that Gal4 or any of the nucleic acid binding protein of Wels et al. needs such stabilization as is taught for the different binding proteins of the '644 patent. The Wels et al. patent discloses generally that nucleic acid binding protein, including Gal4 are satisfactorily functional as presented.

Secondly, the motivation to make the proposed combination does not require a particular placement of a nucleic acid binding domain between a translocation domain and a endoplasmic reticulum retention domain as recited in the base claim. The '644 patent clearly teaches that the benefit of a stabilized disulfide bond can be achieved in other positions of PE than Ib as well. See, for instance, the paragraph bridging cols. 7 and 8 which recites:

> The disulfide-stabilized binding agent may be located at virtually any position within the modified Pseudomonas exotoxin. In one preferred embodiment, the binding agent is inserted in replacement for domain Ia as has been accomplished

in what is known as the TGF.alpha./PE40 molecule (also referred to as TP40) described in Heimbrook et al., Proc. Natl. Acad. Sci., USA, 87: 4697-4701 (1990) and in commonly assigned U.S. Ser. No. 07/865,722 filed Apr. 8, 1992 and in U.S. Ser. No. 07/522,563 filed May 14, 1990.

C. In contravention of MPEP §2143.01, the proposed modification can not render the prior art unsatisfactory for its intended purpose.

As noted by the Examiner, the yeast Gal4 nucleic acid binding domain is a pathogen nucleic acid binding protein which Wels et al. disclose can be used as the nucleic acid binding domain in their nucleic acid transfer system. The cysteine residues of the $Gal4_{(1-149)}$ nucleic acid binding domain form a $Zn(II)_2Cys_6$ binuclear cluster which is critical to the ability of Gal4 to bind its target nucleic acid sequences (see, enclosed, Pan et al., PNAS 87:2077-2081 (1990) particularly the abstract). Altering Gal4 at such would render these cysteines unavailable for binding Zn(II) and hence the Gal4 moiety unsuitable for its intended use.

Alternatively, introduction of additional cysteine molecules into the Gal4 sequence would greatly increase the risk of forming improper disulfide bonds between the cysteine residues of the $Zn(II)_2Cys_6$ binuclear cluster and the extra cysteine. Pastan et al. also teach that reduction in the number of cysteine residues is useful in reducing the opportunities for formation of improper disulfide bonds (see col. 16, lines 41-43).

In view of the above, the Applicants respectfully request that the above rejection be reconsidered and withdrawn.

Response to the Objections to Claims 2, 7, and 8 under 37 CFR §1.75(c)

Without acquiescing to the position of the Examiner and in order to expedite prosecution of the present application, the Applicants have amended the base claim to recite "a polypeptide having an amino acid subsequence." Applicants believe this amendment satisfies the issue raised by the Examiner.

Objection to Claim 2.

As to the objected-to domain III recital, Applicants have amended claim 2 to remove the objectionable recital.

Further comments on the Objection of Claim 2

For the record, Applicants note that Claim 1 recited " translocation domain **comprising** an amino acid subsequence at least 90% identical to the sequence of *Pseudomonas* exotoxin A (PE) (SEQ ID NO:2) from amino acid position 280 to amino acid position 344 thereof and wherein the domain is capable of effecting translocation to the cytosol of the cell." [bold added for emphasis]. The above recital of **comprising** broadens the translocation domain element such that the domain is not limited to the length of an amino acid sequence at least **95%** identical to the sequence of *Pseudomonas* exotoxin A (PE) (SEQ ID NO:2) from amino acid position 280 to amino acid position 344."

Objection to Claim 7.

Applicants believe the amendment to claim 1 also should satisfy the Examiner's concerns with respect to claim 7. In the event that the Examiner's concern was not fully satisfied with the amendment, Applicants note Claim 7 depends from claim 1 and recites "wherein the translocation domain **comprises** the amino acid sequence of SEQ ID NO:2 from the amino acid at position to 280 to the amino acid at position 364." [bolding added for emphasis]. Claim 1 does not require that the translocation domain have a sequence identical to that of SEQ ID NO:2 from the amino acid at position to 280 to the amino acid at position 364. Claim 1 encompasses a 95% degree of identity rather than 100% identity over the recited positions. As claim 7 is narrower than claim 1, Applicants respectfully request that the above objection be reconsidered and withdrawn.

Objection to Claim 8.

Applicants believe the amendment to claim 1 also should satisfy the Examiner's concerns with respect to claim 8. If this is not the case, Applicants further note that Claim 8

depends from claim 1 and recites "wherein the translocation domain is domain II of PE." The traversal of the objection to Claim 2 with respect to the translocation domain subject matter also would applies to claim 8 with equal force. In view thereof, Applicants respectfully request that this objection be reconsidered and withdrawn.

Written Description Rejections Under 35 U.S.C. §112, first paragraph

Claims 1, 46 and 47 stand rejected as to the written description requirement as to requiring mutations or substitutions or variants of a critical subfragment region of domain II. Applicants. Without acquiescing to the position of the Examiner, the Applicants have amended claim 1 to recite **95%** identical. Claim 46 has been canceled. Over a stretch of about 65 amino acids (i.e., PE domain II positions 280-344), 95% identity provides for a substitution or variation of only 3 amino acids.

The Examiner contends that the Applicants have not provided written support for variants within the critical subfragment of Domain II. The Applicants disagree.

The Written Description Guidelines cited by the Examiner provide that:

> What constitutes a ``representative number'' is an inverse function of the skill and knowledge in the art. Satisfactory disclosure of a ``representative number'' depends on whether one of skill in the art would recognize that the applicant was in possession of the necessary common attributes or features of the elements possessed by the members of the genus in view of the species disclosed. In an unpredictable art, adequate written description of a genus which embraces widely variant species cannot be achieved by disclosing only one species within the genus.

(Federal Register 64(244): 71427-71440 Dec. 21, 1999 at middle column 71436).

Here, the art is not so unpredictable and the skill and knowledge of one in the art is quite high. PE is an old protein which has been extensively studied. The specification in the paragraph spanning p.22 and 23 discloses several references already of record which set forth genetically modified forms of PE. Particularly with respect to Domain II, the specification, citing Siegall et al. (1991) *Biochemistry* 30:7154-59, discloses at p. 27, lines 20-24, that amino acids 339 and 343 appear to be necessary for translocation. The enclosed Siegall reference was

incorporated by reference by operation of the last paragraph on p. 62. This reference, co-authored by the instant inventor, also indicates positions which can be mutated with retention of function. For instance, the threonine residue at position 341 can be mutated to a Gln (reference Fig. 4)or a Phe (reference Table II) with retention of function. This reference also teaches that a substitution at the 344 residue (Alanine to Glutamine) retains its activity (reference Table II). The Siegall et al. reference further teaches that additional amino acid substitutions can be made wherein the substituted amino acid is that of a corresponding amino acid from a functionally and structurally similar sequence of Clathrin which has about 29% homology (see paragraph bridging last two columns of the reference). Additionally, the specification teaches the types of conservative amino acid substitutions that can be generally tolerated. Thus, Applicants believe that in view of the state of the art and knowledge available to one of ordinary skill in the art, and particularly in view of the incorporation of the Siegal reference, that one of ordinary skill in the art would consider that the instant inventors were in possession of their invention as based upon the written description at the time of filing.

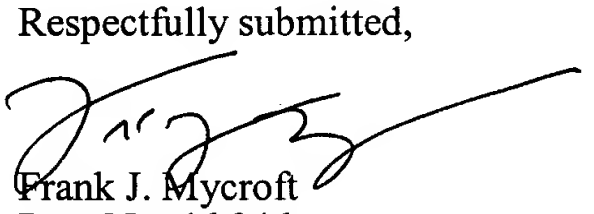
Thus, in view of the above, Applicants request that the instant rejection be reconsidered and withdrawn.

CONCLUSION

In view of the foregoing, Applicants believe all claims now pending in this Application are in condition for allowance and an action to that end is respectfully requested.

If the Examiner believes a telephone conference would expedite prosecution of this application, please telephone the undersigned at 925-472-5000.

Respectfully submitted,



Frank J. Mycroft
Reg. No. 46,946

TOWNSEND and TOWNSEND and CREW LLP
Two Embarcadero Center, Eighth Floor
San Francisco, California 94111-3834
Tel: 925-472-5000
Fax: 415-576-0300
Attachments
FJM:fjm
60354622 v1

Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 3145–3149, May 1989
Biochemistry

Structure and function of the Zn(II) binding site within the DNA-binding domain of the GAL4 transcription factor

TAO PAN AND JOSEPH E. COLEMAN

The Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06510

Communicated by Frederic M. Richards, February 6, 1989 (received for review November 21, 1988)

ABSTRACT **The transcription factor GAL4 from *Saccharomyces cerevisiae* contains a "zinc-finger"-like motif, Cys-Xaa$_2$-Cys-Xaa$_6$-Cys-Xaa$_6$-Cys-Xaa$_2$-Cys-Xaa$_6$-Cys, within its DNA-binding domain. A GAL4 fragment consisting of residues 1–147 plus two additional residues from the cloning vector [denoted GAL4(149*)] has been cloned and overexpressed in *Escherichia coli*. This fragment includes the entire DNA-binding domain (residues 1–74). The homogeneous GAL4-(149*) protein contains 1–1.5 moles of Zn(II) per mole of protein. The GAL4(149*) protein binds tightly to the specific 17-base-pair palindromic DNA sequence found at GAL4 binding sites as shown by gel-retention assays using a ^{32}P-labeled 23-mer containing this sequence. Removal of the intrinsic Zn(II) by EDTA at low pH abolishes binding to the 23-mer. The GAL4(149*) apoprotein can be reconstituted with Zn(II), Cd(II), or Co(II) with restoration of specific DNA binding. Titration of GAL4(149*) apoprotein with ^{113}Cd(II) shows two ^{113}Cd(II) binding sites on the molecule, one with δ of 707 ppm, suggesting coordination to four sulfur atoms, and one with δ of 669 ppm, suggesting coordination to three or four sulfur atoms. Because GAL4(149*) protein contains only six cysteine residues within its DNA-binding domain, the precise coordination of the two Cd(II) ions cannot be stated with certainty; one or more shared $-S^-$ ligands could exist. GAL4(149*) protein contains ≈40% α-helix and ≈20% β-sheet, estimated from circular dichroism. Removal of the native Zn(II) ion causes limited unfolding of secondary structure, but less than one turn of α-helix. The binding of Zn(II), Cd(II), and, to a lesser extent, Co(II) to GAL4(149*) apoprotein protects the protein from proteolysis by trypsin, which produces a 13-kDa DNA-binding core.**

GAL4 is a transcription factor required for galactose utilization in *Saccharomyces cerevisiae* (1). *In vivo*, GAL4 protein binds to 17-base-pair (bp) palindromic sequences (upstream activation sequence G; UAS$_G$) upstream from galactose-inducible genes (2). Although intact GAL4 protein consists of 881 amino acids, the determinants required for recognition of UAS$_G$ sequences are located in the N-terminal 74 amino acids, as shown by deletion mutagenesis (3). The N-terminal 74 amino acids of GAL4 include six cysteine residues, four of which have been predicted to form a single Cys$_2$-Cys$_2$ "zinc finger" (4), analogous to the two Cys$_2$-Cys$_2$ zinc fingers found in the DNA-binding domains of the steroid hormone receptor proteins (5, 6).

To examine the structural basis of this recognition, we have cloned into T7 overexpression vectors in *Escherichia coli* DNA fragments of GAL4 transcription factor consisting of residues (1–74 + 2), (1–92 + 1), and (1–147 + 2). Overproduction of the entire *GAL4* gene product in *E. coli* has not been successful (7 and unpublished results). On the other hand, induction of the engineered genes for all the above fragments resulted in large overproduction of a protein product; however, only the domain consisting of (1–147 + 2) produced a soluble protein. We purified GAL4(1–147 + 2) [denoted GAL4(149*)] in high yield to >95% homogeneity. We found that GAL4(149*) protein isolated from *E. coli* incorporates stoichiometric amounts of Zn(II) and requires Zn(II) for binding to its specific DNA sequence. The native Zn(II) can be replaced by Cd(II) and Co(II) with restoration of specific DNA binding. The characteristics of the GAL4(149*) metalloprotein and apoprotein are reported here.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Cloning and Overproduction of GAL4(149*) Protein. A cDNA encoding GAL4 protein (residues 1–881) was cloned into a *Nde* I/*Hin*dIII site of a modified pAR3039 (ref. 8; unpublished results) by ligating the *Sph* I/*Hin*d III fragment from pBM1123 (supplied by M. Johnston, Washington University, Saint Louis) to a synthetic Nde I/*Sph* I linker. A *Spe* I stop-codon linker was then introduced into the *Cla* I site at amino acid 147 to yield pTPT7G1 with GAL4(147) protein under the control of the T7 RNA polymerase promoter. The *Spe* I linker adds two amino acids (Leu-Asp) onto GAL4(1–147) protein. The natural sequence after amino acid 147 is Ile-148 and Asp-149. Hence, we have termed our construct GAL4(149*). Overproduction of GAL4(149*) construct was achieved by adding isopropyl β-D-thiogalactose (1 mM) to BL21(DE3) cells/pTPT7G1 at mid-logarithmic phase, and the cells were harvested 4 hr later. GAL4(149*) protein represented ≈10% of the total *E. coli* protein.

Purification of GAL4(149*) Construct. The purification procedure reported for GAL4 fragments (7) did not yield high amounts of GAL4(149*). The procedure to be described results in much higher yields, ≈2 mg of GAL4(149*) protein per g (wet weight) of cells. Fifteen grams of cells were suspended in 40 ml of lysis buffer (50 mM Tris·HCl, pH 8.0/200 mM KCl/1 mM EDTA/1 mM dithiothreitol) and sonicated twice for 3 min on ice. Polymin P was then added over a 10-min period to 0.5%, followed by centrifugation at 25,000 × *g* for 20 min at 4°C. After dialysis of the supernatant against standard column buffer (STD) [10 mM Tris·HCl, pH 8.0/1 mM EDTA/1 mM 2-mercaptoethanol/10% (vol/vol) glycerol] plus 50 mM NaCl, the dialysate was loaded onto a pre-equilibrated Trisacryl-SP column and washed sequentially with STD plus 50 mM NaCl, STD plus 150 mM NaCl, and STD plus 300 mM NaCl. The fractions from the 300 mM NaCl wash were combined and dialyzed against STD plus 50 mM NaCl. The dialysate was then loaded onto a Cibacron Blue 3GA agarose column, washed with STD plus 50 mM NaCl and STD plus 250 mM NaCl, and then eluted with a NaCl gradient, STD plus 250–1000 mM NaCl. GAL4(149*) protein elutes in the 500–1000 mM NaCl fractions and is >95% pure.

Gel-Retardation Assays. Gel-retardation assays to detect GAL4(149*) binding to DNA were performed on 6% polyacrylamide gels as described (9). The buffer was 20 mM Tris·HCl/1 mM dithiothreitol/80 mM NaCl, pH 8.0 (buffer G), and sample volume was 20 μl. *GAL4*-specific DNA was

Table 1. Metal content of GAL4(149*) construct

GAL4(149*)	Treatment	Mol/mol of protein Zn	Cd	Co
Native	Dialysis (prep. 1)	1.27 ± 0.02	–	–
	Dialysis (prep. 2)	1.49 ± 0.02	–	–
	Dialysis (prep. 3)	1.14 ± 0.03	–	–
Native	Dialysis + 10 mM EDTA, pH 8	1.01 ± 0.04	–	–
Apoprotein	Dialysis + 10 mM EDTA, pH 5	0.05 ± 0.01	–	–
Cd(II)†	Dialysis + 1 mM $CdCl_2$,‡ pH 8	0.30	0.68	–
Co(II)	Dialysis + 100 mM $CoCl_2$,§ pH 5	0.05	–	1.15¶

†Cd(II)GAL4(149*) metalloprotein can also be formed by addition of stoichiometric amount of metal ion to the apoprotein.
‡Dialysis was done against 1 mM $CdCl_2$ at 4°C for 20 hr, followed by dialysis versus metal free buffer.
§Dialysis was carried out twice against 100 mM $CoCl_2$ at 4°C for 20 hr under nitrogen, followed by dialysis versus metal free buffer, pH 8.
¶Co(II) can be removed by exhaustive dialysis.

a synthetic ^{32}P-labeled 23-bp oligonucleotide (sequence shown below), incorporating a single *GAL4* 17-bp recognition sequence (3).

```
5'-GATCC CGGAAGACTCTCCTCCG G
       G GCCTTCTGAGAGGAGGC CCTAG
```

^{113}Cd NMR. The NMR was performed on a Bruker AM-500 spectrometer (110.93 MHz for ^{113}Cd) with a 10-mm broadband probe. Samples were ≈0.54 mM GAL4(149*) protein/50 mM phosphate/250 mM NaCl, pH 8.0 at 25°C; sample volumes were 2.0 ml. A 45° pulse and a recycle time of 2 sec were used.

Zinc, Cadmium, and Cobalt Analyses. These were performed by atomic absorption spectroscopy using an Instrumentation Laboratory (Lexington, MA) IL157 spectrometer.

RESULTS

Metal Content of GAL4(149*) Construct. GAL4(149*) contains from 1 to 1.5 mol of Zn(II) per mole of protein (Table 1). Zn(II) cannot be removed by EDTA at pH 8, but exhaustive dialysis will reduce the Zn(II) content to 1 mol per mol of protein. GAL4(149*) apoprotein can be prepared by dialysis at pH 5 against 10 mM EDTA/50 mM sodium acetate buffer. After removal of the Zn(II)–EDTA complex by extensive dialysis, the pH can be returned to 8 with the production of a stable apoprotein that contains 0.05 mol of Zn(II) per mol. Zn(II) can be reconstituted to the apoprotein by adding a 1:1 molar ratio of the metal ion in buffer G and incubating for 60 min at 4°C. Both Cd(II) and Co(II) can be substituted for Zn(II) (Table 1).

Binding of GAL4(149*) to DNA. Native GAL4(149*) retains the ^{32}P-labeled 23-mer double-stranded DNA incorporating the 17-bp-specific sequence on the gel-retention assay (Fig. 1*A*). When the apoprotein is used, there is no retention of the DNA. Two complexes between GAL4(149*) protein and the specific 23-mer, a major and a minor one, which migrate at slightly different rates are seen on the gel-retention assay. The great specificity of the Zn(II)-dependent binding of GAL4 protein to the upstream activation sequence G (UAS_G) DNA is best demonstrated by the gel-retention assay, in which there is a large excess of both competing nonspecific DNA as well as GAL4 protein (Fig. 1*A*).

Readdition of Zn(II) in a 1:1 molar ratio to the apoprotein reestablishes retention of the specific 23-bp fragment (Fig. 1*A*). There may be a slight increase in the efficiency of retention at the 2:1 and 3:1 molar ratios of Zn(II) to protein. Increasing the Zn(II)-to-protein ratio to 50:1 has no further effect. Cd(II) can be substituted for Zn(II) with full retention of the specific DNA at a 1:1 Cd(II)-to-protein ratio. Co(II) will induce some binding of the specific DNA, but a 5:1

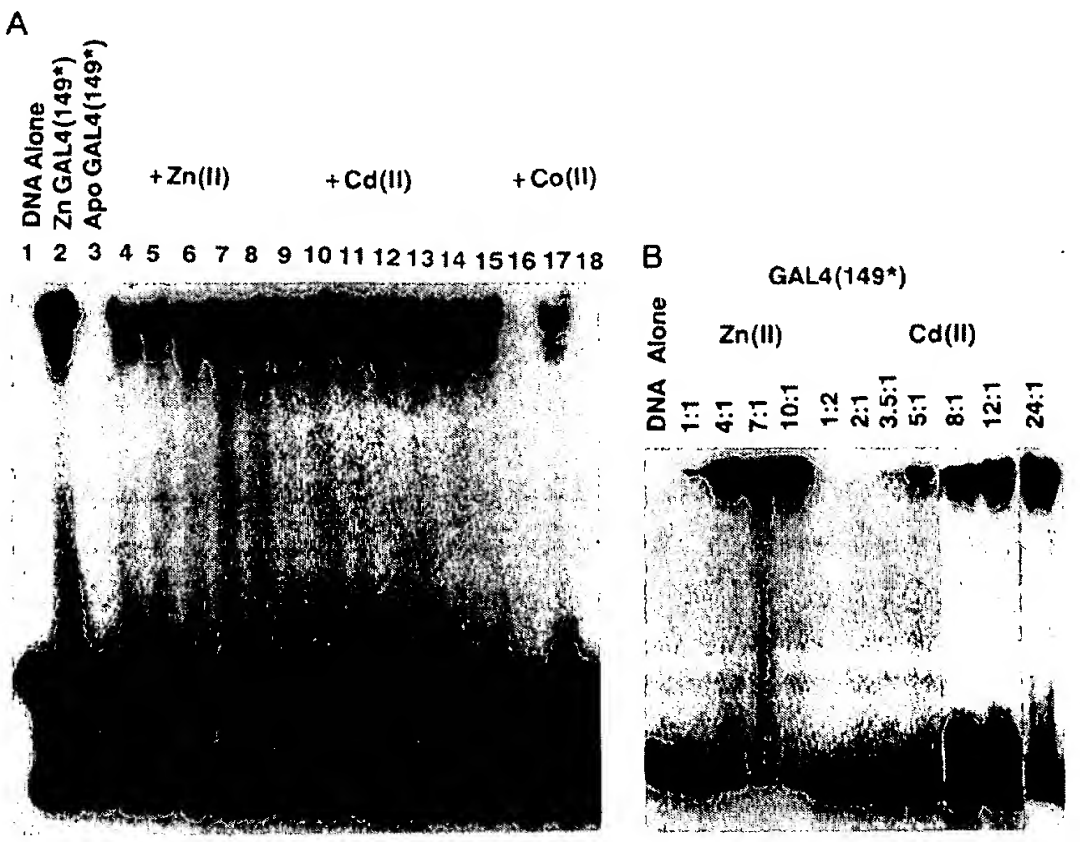


FIG. 1. Gel retardation of DNA by GAL4(149*) protein. (*A*) Retention of *GAL4*-specific ^{32}P-labeled 23-mer (10 nM) by 1 μM of purified GAL4(149*) protein in the presence of 20 μM unlabeled calf thymus DNA. Lanes: 1, DNA alone; 2, Zn(II)GAL4(149*); 3, GAL4(149*) apoprotein; 4–9, GAL4(149*) apoprotein reconstituted with 1:1, 1:2, 1:3, 1:5, 1:10, and 1:50 protein-to-Zn(II) ratios; 10–15, GAL4(149*) apoprotein reconstituted 1:1, 1:2, 1:3, 1:5, 1:10, and 1:50 protein-to-Cd(II) ratios; 16–18, GAL4(149*) apoprotein reconstituted with 1:2, 1:5, and 1:50 protein-to-Co(II) ratios. (*B*) Retention of GAL4-specific ^{32}P-labeled 23-mer (0.1 μM) as a function of the protein-to-DNA ratio of native Zn(II)GAL4(149*) and reconstituted Cd(II)GAL4(149*). The reaction mix contained 1 μM calf thymus DNA. Ratios above each lane indicate the protein-to-DNA ratio.

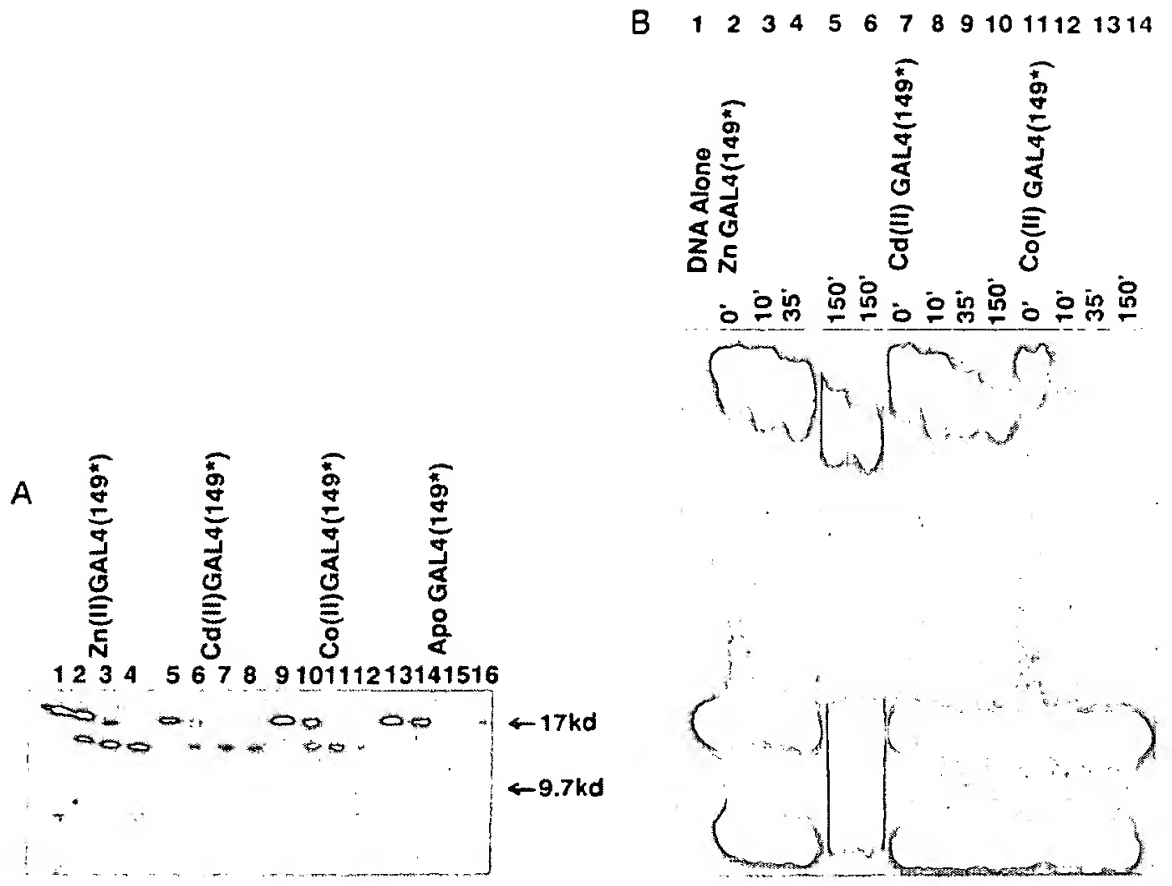


FIG. 2. Limited proteolysis of GAL4(149°) protein by trypsin. (*A*) Time course of trypsin proteolysis. Samples were withdrawn at 0, 10, 35, and 150 min. Trypsin cleavage was stopped by boiling in SDS buffer for 5 min. Lanes: 1–4, Zn(II)GAL4(149°); 5–8, Cd(II)GAL4(149°); 9–12, Co(II)GAL4(149°); 13–16, GAL4(149°) apoprotein. (*B*) Gel retardation of ^{32}P-labeled *GAL4*-specific 23-mer by the tryptic fragments of GAL4(149°) protein. Trypsin cleavage was stopped by addition of soybean trypsin inhibitor in molar excess to trypsin before mixing with the ^{32}P-labeled 23-mer. Samples were withdrawn at 0, 10, 35, and 150 min after starting hydrolysis. Lanes: 1, DNA alone; 2–5, Zn(II)GAL4(149°); 6, Zn(II)GAL4(149°), 150 min at a GAL4(149°)-to-trypsin ratio of 100:1 (wt/wt); 7–10, Cd(II)GAL4(149°); 11–14, Co(II)GAL4(149°). In both *A* and *B* 10 μM GAL4(149°) protein was mixed with a 400:1 (wt/wt) ratio of protein to trypsin (unless otherwise noted) in buffer G at 25°C. Co(II)GAL4(149°) contained 1.15 mol of Co(II) per mol of protein.

Co(II)-to-protein molar ratio is required before significant binding of the specific DNA is seen. Excess Co(II) destroys the binding to DNA.

The gel-retardation assay does not lend itself to the determination of the minimum protein-to-DNA stoichiometry present in a given complex. While it is sometimes possible to isolate a 1:1 protein–DNA complex by cutting out the complex from the gel, the several systems we have investigated by this technique require an excess of protein at the start to retain most DNA with the complex. This effect appears to derive from the loss of free protein as the sample enters the gel, because free protein migrates either toward the cathode or not at all. To better compare the affinities of native Zn(II)GAL4(149°) and the reconstituted Cd(II)GAL4(149°) for the specific 23-mer, gel-retention assays for both proteins as a function of the initial protein-to-DNA ratio in the presence of a 10:1 excess of calf thymus DNA are shown in Fig. 1*B*. The native Zn(II) protein requires a 10:1 molar ratio of protein to DNA to retain all labeled DNA. The Cd(II) protein requires approximately twice the concentration of the Zn(II) protein to retain comparable amounts of DNA (Fig. 1*B*). A moderate drop in DNA binding affinity has previously been seen when Cd(II) replaces Zn(II) in the single-stranded DNA binding protein, gene 32 protein from T4 (10). The GAL4(149°) apoprotein retains no DNA under these conditions. Zn(II)GAL4(149°) demonstrates some binding to a ^{32}P-labeled DNA of random sequence in the absence of calf thymus DNA (data not shown). The random DNA, however, is completely displaced by calf thymus DNA; this nonspecific binding mode is also Zn(II) dependent.

Zn(II) Protects GAL4(149°) Against Proteolysis by Trypsin. When the native Zn(II)GAL4(149°) is treated with trypsin [1:400 (wt/wt) trypsin to GAL4(149°)], rapid but limited proteolysis results in a 13-kDa core that resists further cleavage (Fig. 2*A*). In contrast, exposure of the GAL4(149°) apoprotein to trypsin rapidly degrades the protein to small peptides. After reconstitution with Cd(II) limited proteolysis is once again seen. Co(II) is less effective than either Zn(II) or Cd(II) in restoring resistance to complete proteolysis, and the 13-kDa Co(II) protein core is rapidly cleaved by trypsin. The 13-kDa core can bind to the specific recognition sequence for GAL4 (Fig. 2*B*).

Circular Dichroism of Zn(II)GAL4(149°) and GAL4(149°) Apoprotein. The native Zn(II)GAL4(149°) has significant molar ellipticity in the wavelength region of the peptide bond chromophores, -23.0×10^5°·cm^2 per dmol at 208 nm and -15.8×10^5°·cm^2 per dmol at 222 nm (Fig. 3). These values require that considerable amounts of α-helical and β-sheet structure be present in GAL4(149°) protein. Although the CD spectrum cannot be precisely matched by a combination of the three CD curves reported for homopolypeptides in the α, β, and random-coil configurations, a reasonable graphical fit is obtained by a combination of 40% α-helix, 20% β-sheet, and 40% random coil (11). Application of a computer program for calculating the Chou–Fasman prediction of secondary structure including β turns from the amino acid sequence gives 50% α-helix, 20% β-sheet, and 30% random coil for the GAL4(149°) amino acid sequence. Removal of Zn(II) results in a relatively small but significant change in the secondary structure, as shown by a decrease in negative molar ellipticity of $+7.5 \times 10^4$° · cm^2 per dmol at 222 nm and $+3 \times 10^5$°·cm^2 per dmol at 208 nm. Reconstituted Cd(II)GAL4(149°) shows the same CD spectrum as the native protein.

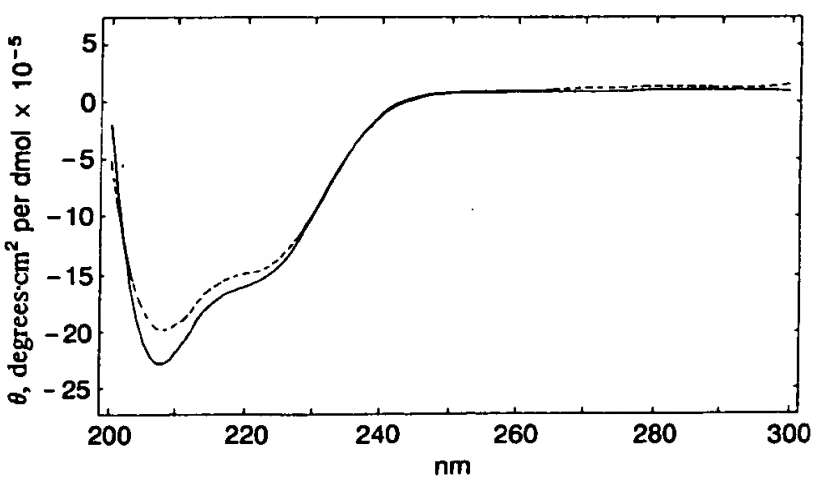


FIG. 3. Circular dichroism of Zn(II)GAL4(149°) and GAL4(149°) apoprotein. CD spectra of both Zn(II)GAL4(149°) (—) and GAL4(149°) apoprotein (----) in 10 mM Tris/150 mM NaCl/1 mM 2-mercaptoethanol/50 μM EDTA/10% glycerol, pH 8.0 at 28°C. Protein concentrations were 5 μM for both species.

Spectroscopic Properties of the Co(II)GAL4(149*) and Cd(II)GAL4(149*). Co(II) optical spectra: Visible d–d transitions of Co(II) can be used to distinguish octahedral from tetrahedral ligand geometry around Co(II), and near-UV sulfur to Co(II) charge-transfer bands can often be used to detect sulfur ligands to the Co(II). Although the Co(II)-substituted GAL4(149*) protein has the intense charge-transfer bands at 340 nm ($\varepsilon = 2.2 \times 10^3$ $mol^{-1} \cdot cm^{-1}$) and 305 ($\varepsilon = 3 \times 10^3$ $mol^{-1} \cdot cm^{-1}$) expected from sulfur ligation, there are no intense d–d transitions in the vicinity of 700 nm typical of tetrahedral Co(II) complexes. A comparison of the absorption spectrum of Co(II)GAL4(149*) protein to the tetrahedral four $-S^-$ ligand site in Co(II) alcohol dehydrogenase (12) is shown in Fig. 4. A Co(II) protein spectrum similar to this, with no tetrahedral Co(II) d–d transitions, is observed for the Co(II)-substituted B site or exchangeable Zn(II) site in *E. coli* RNA polymerase (13).

^{113}Cd NMR of ^{113}Cd(II)GAL4(149*) Protein. When two equivalents of ^{113}Cd(II) are added to GAL4(149*) apoprotein, the ^{113}Cd(II) NMR shows two signals of equal intensity at δ of 707 and 669 ppm (Fig. 5). The chemical shifts of ^{113}Cd(II) bound to protein sites with known ligand composition show that ^{113}Cd chemical shifts can be used to predict the number of sulfur donor atoms coordinating the Cd(II) ion (14–16). The δ of 707 ppm is within the chemical shift range expected for a S_4 donor set, whereas a δ of 669 ppm is compatible with either a S_3X or S_4 set (see *Discussion*). At present we have been unsuccessful in attempting to generate a GAL4(149*) protein with only one of these sites occupied with ^{113}Cd(II), either by addition of one ^{113}Cd(II) or removal of one of the bound ^{113}Cd(II) ions after reconstitution. Both ^{113}Cd(II) ions seem to be removed simultaneously, suggesting there may be some cooperative binding.

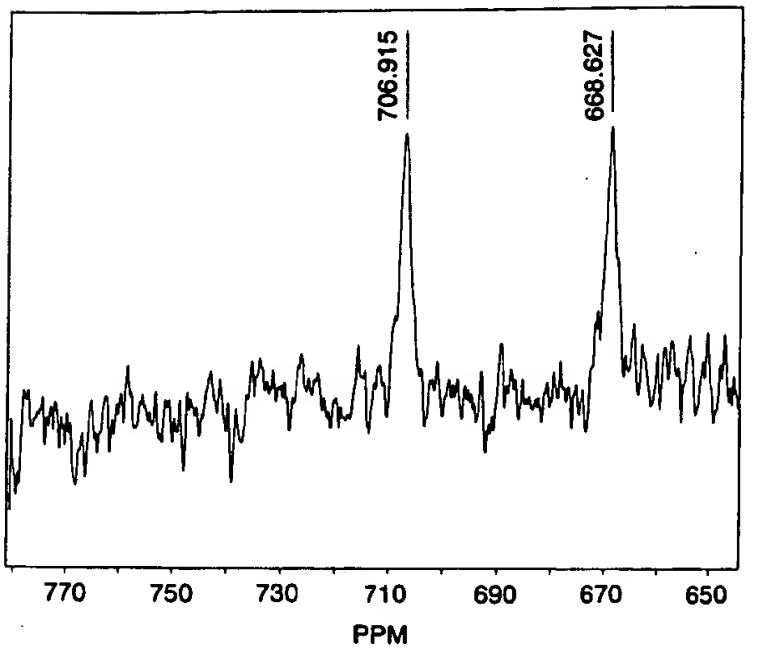


FIG. 5. ^{113}Cd NMR spectrum of GAL4(149*) protein. Spectral width is 15.2 kHz (136 ppm), and the number of transients is 9800. A 50-Hz line broadening was applied for spectral enhancement; acquisition conditions are described in text. Chemical shift is plotted relative to that of 0.1 M $^{113}CdClO_4$, $\delta = 0$. PPM, parts per million.

DISCUSSION

The zinc analysis on GAL4(149*) protein demonstrates that GAL4 is a zinc metalloprotein (Table 1). Likewise the recognition of the specific DNA sequence to which GAL4 transcription factor binds depends on the presence of Zn(II) (Fig. 1*A*). The sequence of the GAL4(149*) fragment is given in Fig. 6. Six cysteine residues occur in this fragment, Cys-11, Cys-14, Cys-28, and Cys-31, proposed to be involved in formation of a Zn(II) complex (4), and two additional cysteine residues, Cys-21 and Cys-38. Mutation studies on GAL4 (18) revealed that two regions within the DNA binding domain (residues 1–74) are important in specific DNA binding. All missense mutations bearing a single amino acid change in the immediate vicinity of the putative zinc-finger (residues 10–38) abolish specific DNA binding, whereas most missense mutations in the region downstream (residues 40–50) are accompanied by partial loss of specific DNA-binding affinity.

Cd(II) effectively substitutes for Zn(II) in restoring the specific DNA-binding function to GAL4(149*) apoprotein (Fig. 1). On the other hand, Co(II) is not as effective in restoring the binding of GAL4(149*) to its specific DNA sequence (Fig. 2*B*). The native GAL4(149*) protein assumes a significant amount of α-helical and β-sheet secondary structure as shown by the UV CD (Fig. 3). The removal of the Zn(II) must result in only modest unfolding of this secondary structure, because the molar ellipticity becomes less negative by only $7.5 \times 10^{4\circ} \cdot cm^2$ per dmol at 222 nm and $3 \times 10^{5\circ} \cdot cm^2$ per dmol at 208 nm (Fig. 3). When the change at 222 nm is

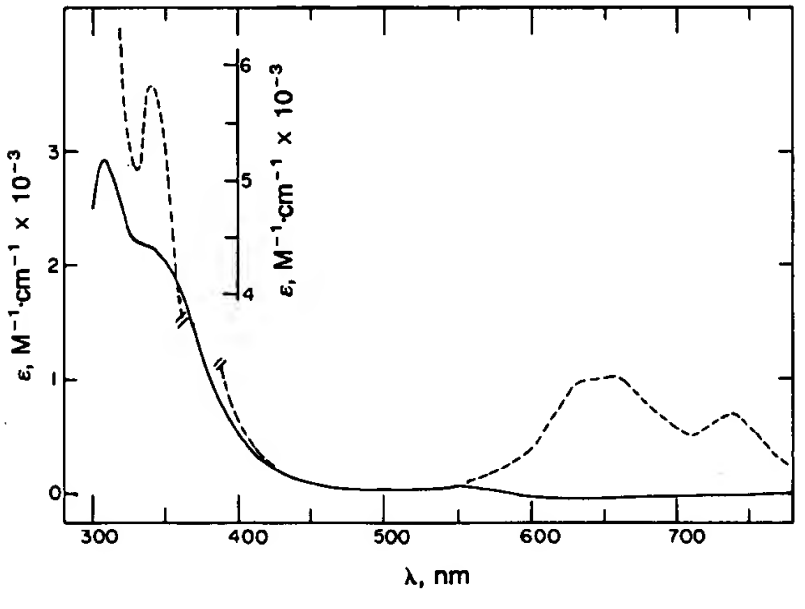


FIG. 4. Visible absorption spectrum of Co(II)GAL4(149*) protein (—). The spectrum is corrected against Zn(II)GAL4(149*) of the same *A* at 280 nm. The concentration of Co(II)GAL4(149*) used for this spectrum is ≈40 μM. (---) Horse liver alcohol dehydrogenase with Co(II) substituted at the structural site, replotted from ref. 12. The extinction coefficient between 300–350 nm for liver alcohol dehydrogenase should be referred to the middle ordinate.

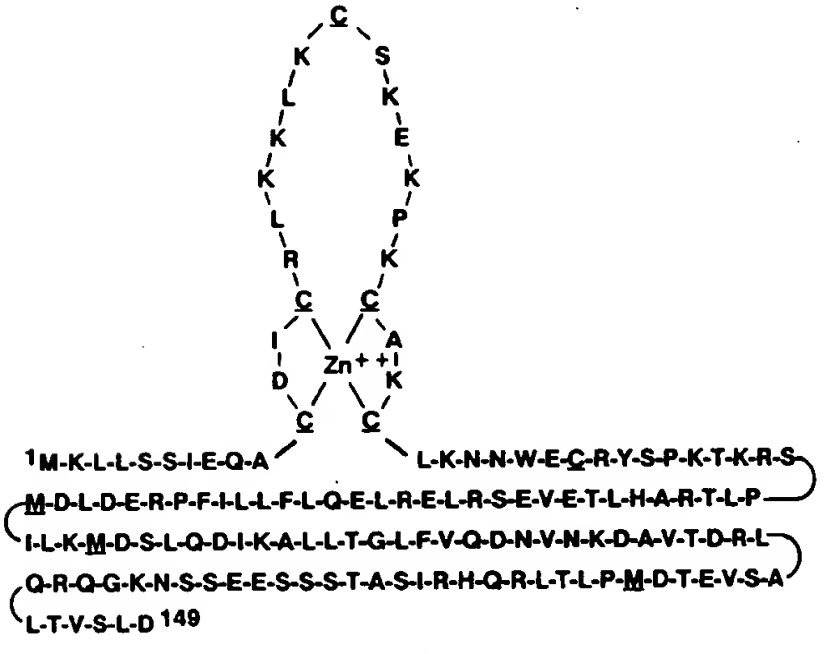


FIG. 6. Amino acid sequence of the GAL4(149*) protein. The amino acid sequence (one-letter code) of the intact GAL4 transcription factor is taken from ref. 17, and the proposed Zn(II) ligands are as indicated in ref. 4.

assumed to be unfolding of α-helix, then only $\approx$ two residues are induced to form α-helix by Zn(II) binding. The considerably larger change in ellipticity at 208 nm suggests that unfolding of some other type of secondary structure also occurs on Zn(II) removal, quite possibly β-turns. It seems reasonable to conclude that Zn(II) binding is involved in inducing the final conformation of a section of the polypeptide backbone of GAL4(149*) protein involved directly or indirectly in forming the DNA-binding surface. Disorganization of this section on Zn(II) removal must account for the loss of specific DNA recognition (Fig. 1*A*).

Extensive proteolysis and differential scanning calorimetry studies of the Zn(II) binding domain in the single-stranded DNA-binding protein, gene 32 protein from bacteriophage T4, show that Zn(II) protects a core domain from further proteolysis as well as imparting a significant increase in thermal stability to the protein (10, 19). Likewise for GAL4-(149*) protein, the Zn(II) protects a 13 kDa core from complete proteolysis (Fig. 2). Because this core contains the DNA-binding surface (Fig. 2*B*), it probably represents the minimal DNA-binding subdomain.

There are two general types of zinc-finger domains in transcription factors: those with the ligand composition, Cys_2-His_2—e.g., transcription factor IIIA from *Xenopus* oocytes (20, 21)—and those with the ligand composition, Cys_2-Cys_2—e.g., the steroid receptor proteins (5, 6). Few physicochemical studies have been reported on representatives of the Cys_2-Cys_2 group, but both the DNA-binding fragment of the glucocorticoid receptor and now the appropriate GAL4 fragment contain the expected Zn(II) content (Table 1) (6). The structure surrounding the metal ion is not clear in proteins belonging to this group. Number and types of amino acid residues between pairs of putative cysteine ligands is highly variable, and the sequence often contains extra cysteines. Hence, precise folding schemes have not been well defined. For GAL4(149*), the present data show that, while the metal ion induces additional folding of the polypeptide chain, the induced folding must be highly localized within a well-defined secondary structure and does not appear to increase the content of α-helix (Fig. 3).

The optical absorption spectrum of the Co(II) derivative of GAL4(149*) protein does not suggest tetrahedral geometry (Fig. 4). Co(II) in this case, however, may not induce the normal configuration of the coordination complex, because the restoration of DNA binding is clearly defective (Fig. 1*A*). ^{113}Cd NMR of the ^{113}Cd(II) GAL4(149*) protein shows that a metal site with four S^- ligands exists in GAL4(149*) protein, the site with δ of 707 ppm (Fig. 5). Most such sites in proteins have been found to consist of some variant of tetrahedral coordination geometry. The human glucocorticoid receptor has been shown to contain two Zn(II) ions per molecule within the DNA-binding domain of 150 amino acid residues that contains two Cys_2-Cys_2 zinc fingers (5, 6). The presence of the two Zn(II) ions is required for the binding of the 150-residue fragment to the specific DNA sequence recognized by the glucocorticoid receptor. Extended x-ray absorption fine structure of the Zn(II) protein fragment suggests that both Zn(II) ions occupy tetrahedral sites consisting of four S^- ligands (6).

The unexpected finding in the case of GAL4(149*) protein is the presence of a second metal binding site with coordination to at least three sulfur donors (Fig. 5). At present we cannot state whether the occupancy of this site affects DNA binding. Although GAL4(149*) protein, as isolated, consistently contains >1 mol of Zn(II) per mol, the protein can be treated with metal-free solutions, such that the average Zn(II) content is reduced to one mol per mol without obvious loss of DNA binding. The conditions of the gel-retention assay, however, do not lend themselves to careful correlation of DNA-binding affinity as a function of Zn(II) stoichiometry. In fact, the gel in Fig. 1*A* could be interpreted to show some difference in efficiency of retention between the 1:1, 2:1, and 3:1 Zn(II) stoichiometries, although this is less evident for the Cd(II) derivative.

The distribution of ligands in GAL4(149*) protein that provide for sites with two such downfield ^{113}Cd NMR signals is not obvious in the absence of more structural information. Although four $-S^-$ coordination can explain one of the ^{113}Cd NMR signals, there are only two additional $-S^-$ ligands possible. Possibly the two Cd(II) ions share one or more $-S^-$ ligands, creating a two-metal cluster. Lack of ^{113}Cd–^{113}Cd coupling might be thought to rule this alternative out, but such coupling (30–50 Hz) may not be easy to resolve. There are also three methionine residues toward the C-terminal region of GAL4(149*) transcription factor, and the sulfur in this ether linkage can be a donor to Cd(II) (22). The methionine residues are not within the sequence involved in DNA binding, but the second Cd(II) site may not directly affect ligand binding. More definitive definition of structure of the metal complexes awaits more precise probes like extended x-ray absorption fine structure and the examination of fluorescence quenching upon nucleotide binding to gain a more exact measure of DNA-binding affinity versus metal ion stoichiometry.

We thank Mark Johnston for the gift of the plasmid containing the cDNA for GAL4 protein and David Giedroc for many helpful suggestions. This work was supported by National Institutes of Health Grants DK09070 and GM21919. The 500-MHz NMR was supported by National Institutes of Health Grant RR03475, National Science Foundation Grant DMB-8610557 and American Cancer Society Grant RD259. This work is in partial fulfillment of the requirements for the Ph.D. degree (T.P.).

1. Oshima, Y. (1982) in *Molecular Biology of the Yeast Saccharomyces*, eds. Strathern, J., Jones, E. & Broach, J. K. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), Vol. 1, pp. 159–180.
2. Giniger, E., Varnum, S. M. & Ptashne, M. (1985) *Cell* 40, 767–774.
3. Keegan, L., Gill, G. & Ptashne, M. (1986) *Science* 231, 699–704.
4. Johnston, M. (1987) *Nature (London)* 328, 353–355.
5. Severne, Y., Wieland, S., Schaffner, W. & Rusconi, S. (1988) *EMBO J.* 7, 2503–2508.
6. Freedman, L. P., Luisi, B. F., Korszun, Z. R., Basavappa, R., Sigler, P. B. & Yamamoto, K. R. (1988) *Nature (London)* 334, 543–546.
7. Lin, Y.-S., Carey, M. F., Ptashne, M. & Green, M. R. (1988) *Cell* 54, 659–664.
8. Studier, F. W. & Moffatt, N. (1986) *J. Mol. Biol.* 189, 113–130.
9. Fried, M. G. & Crothers, D. M. (1981) *Nucleic Acids Res.* 5, 6505–6525.
10. Giedroc, D. P., Keating, K. M., Williams, K. R. & Coleman, J. E. (1987) *Biochemistry* 26, 5251–5259.
11. Greenfield, N. & Fasman, G. D. (1969) *Biochemistry* 8, 4108–4116.
12. Maret, W., Andersson, I., Dietrich, H., Schneider-Bernlohr, H., Einarsson, R. & Zeppezauer, M. (1979) *Eur. J. Biochem.* 98, 501–509.
13. Giedroc, D. P. & Coleman, J. E. (1986) *Biochemistry* 25, 4969–4978.
14. Bobsein, B. R. & Myers, R. J. (1980) *J. Am. Chem. Soc.* 102, 2454–2455.
15. Armitage, I. M. & Otvos, J. D. (1982) in *Biological Magnetic Resonance*, eds. Berliner, L. J. & Reuben, J. (Plenum, New York), Vol. 4, pp. 79–144.
16. Giedroc, D. P., Johnson, B. A., Armitage, I. M. & Coleman, J. E. (1989) *Biochemistry* 28, 2410–2418.
17. Laughon, A. & Gesteland, R. F. (1984) *Mol. Cell. Biol.* 4, 260–267.
18. Johnston, M. & Dover, J. (1987) *Proc. Natl. Acad. Sci. USA* 84, 2401–2405.
19. Keating, K. M., Ghosaini, L. R., Giedroc, D. P., Williams, K. R., Coleman, J. E. & Sturtevant, J. M. (1988) *Biochemistry* 27, 5240–5245.
20. Miller, J., Mclachlan, A. P. & Klug, A. (1985) *EMBO J.* 4, 1609–1614.
21. Diakun, G. P., Fairall, C. & Klug, A. (1986) *Nature (London)* 324, 698–699.
22. Engeseth, H. R., McMillin, D. & Otvos, J. D. (1984) *J. Biol. Chem.* 259, 4822–4826.

hyperdictionary

English Dictionary Computer Dictionary Thesaurus Dream Dictionary Medical Dictionary

Search Dictionary: cer Search

Meaning of CER

WordNet Dictionary

Definition: [n] an emotional response that has been acquired by conditioning

Websites:

- **Find the Best Sites For Cer With Starware**
  Starware search is an excellent resource for quality sites on Cer and much more! Starware also provides related listings for Cer.
  search.starware.com

Synonyms: conditioned emotion, conditioned emotional response

See Also: emotion

Computing Dictionary

Definition: Canonical Encoding Rules

Analysis of Sequences in Domain II of *Pseudomonas* Exotoxin A Which Mediate Translocation

Clay B. Siegall,‡ Masato Ogata, Ira Pastan, and David J. FitzGerald*

Laboratory of Molecular Biology, Division of Cancer Biology Diagnosis and Centers, National Cancer Institute, National Institutes of Health, Building 37, Room 4E16, Bethesda, Maryland 20892

Received December 26, 1990; Revised Manuscript Received April 16, 1991

ABSTRACT: *Pseudomonas* exotoxin (PE) contains 613 amino acids that are arranged into 3 structural domains. PE exerts its cell-killing effects in a series of steps initiated by binding to the cell surface and internalization into endocytic vesicles. The toxin is then cleaved within domain II near arginine-279, generating a C-terminal 37-kDa fragment that is translocated into the cytosol where it ADP-ribosylates elongation factor 2 and arrests protein synthesis. In this study, we have focused on the functions of PE which are encoded by domain II. We have used the chimeric toxin TGFα-PE40 to deliver the toxin's ADP-ribosylating activity to the cell cytosol. Deletion analysis revealed that sequences from 253 to 345 were essential for toxicity but sequences from 346 to 364 were dispensable. Additional point mutants were constructed which identified amino acids 339 and 343 as important residues while amino acids 344 and 345 could be altered without loss of cytotoxic activity. Our data support the idea that domain II functions by first allowing PE to be processed to a 37-kDa fragment and then key sequences such as those identified in this study mediate the translocation of ADP-ribosylation activity to the cytosol.

Protein toxins are found in many organisms and function as potent cytotoxic agents. Many of these toxins kill cells by attacking the cellular machinery which is used to synthesize proteins (Olsnes & Sandvig, 1988). Directing the action of toxins to specific cells can be accomplished by replacing the native cell recognition region of the toxin with a ligand that specifically binds only to certain cells. There are many examples of targeted cell-killing using protein toxins such as *Pseudomonas* exotoxin, diphtheria toxin, and ricin (Pastan & FitzGerald, 1989; Pastan et al., 1986; Vitetta et al., 1987; Estworthy & Neville, 1984; Greenfield et al., 1987; FitzGerald & Pastan, 1989). We have been employing *Pseudomonas* exotoxin A (PE) to kill a variety of cell types by attaching PE or modified forms of PE to different ligands (Chaudhary et al., 1987, 1988; Siegall et al., 1988, 1989b; Lorberboum-Galski et al., 1988; Ogata et al., 1989; Batra et al., 1989). Along with their utility in killing target cells, chimeric toxins can be used to identify important sequences of toxin function (Johnson & Youle, 1989; Siegall et al., 1989a; Williams et al., 1990; Chaudhary et al., 1990).

PE is composed of three structural domains as determined by crystallographic analysis (Allured et al., 1986). Domain I is made up of two subunits, domain Ia (amino acids 1–252) and domain Ib (amino acids 365–404), which are closely associated in the crystal structure but are separated by domain II in the amino acid sequence (Gray et al., 1984). The amino acids of domain Ia encode the PE cell recognition function (Hwang et al., 1987), while the function of domain Ib is not yet known (Siegall et al., 1989a). Mutational analysis of the lysine residues in domain Ia indicated that Lys-57 is essential for cell binding (Jinno et al., 1988). Further, mutational analysis revealed that amino acids 246, 247, and 249 are also important for killing by the native toxin.

Domain II (amino acids (253–364), which contains six consecutive α-helices, has been demonstrated to be responsible for translocation of the toxin across membranes (Hwang et al., 1987; Siegall et al., 1989a; Jinno et al., 1989; Edwards et al., 1989). Previous studies have determined that there are several distinct regions within domain II which are essential for toxin activity. Processing of PE to its active form requires the amino end of domain II, with the actual cleavage occurring around amino acid 279 (Ogata et al., 1990). Deletion of the amino end of domain II, amino acids 254–263, inactivates the cytotoxicity of PE, but deletion of the carboxyl end (361–364) and part of domain Ib (amino acids 361–380) does not affect the cytotoxicity of PE (Siegall et al., 1989a). When amino acids 337–380 were deleted, PE cytotoxicity was lost, indicating that there was an essential region between aa 337 and 360 (Siegall et al., 1989a). We have also determined the importance of several arginine residues in domain II (Jinno et al., 1989). The results indicate that the arginines at aa 276, 279, 330, and 337 are important for full PE toxicity. In particular, mutations at Arg-276 or Arg-279, which lie in a loop on the surface of domain II, completely inactivated the toxin. We have recently determined that cellular proteolytic processing of PE required for translocation of a 37-kDa fragment of PE occurs within this loop (Ogata et al., 1990).

Domain III, which functions as the enzymatic unit of the toxin contains the ADP-ribosylation activity. PE kills cells by ADP-ribosylating elongation factor 2 and halting protein synthesis. The structural boundary of domain III as defined by crystallography structure is 405–613 (Allured et al., 1986). The functional NH_2-terminal boundary of domain III is amino acid 400, and the carboxyl boundary is 601 (Siegall et al., 1989a; Chaudhary et al., 1990). Previous studies of domain III have been shown that His-426, Tyr-481, Glu-553, and Trp-558 are essential for ADP-ribosylation (Carroll & Collier, 1987; Wozniak et al., 1988; Brandhuber et al., 1988; Lukas & Collier, 1988) and the carboxyl five amino acids (609–613) are important for translocation of PE into the cytosol (Chaudhary et al., 1990).

In the current study, we have focused on the role of amino acids at the carboxyl end of domain II containing the E helix (aa 333–350) and F helix (aa 361–364) in the cytotoxic action

* To whom correspondence should be addressed.
‡ Present address: Pharmaceutical Research Institute, Bristol-Myers Squibb, Wallingford, CT.

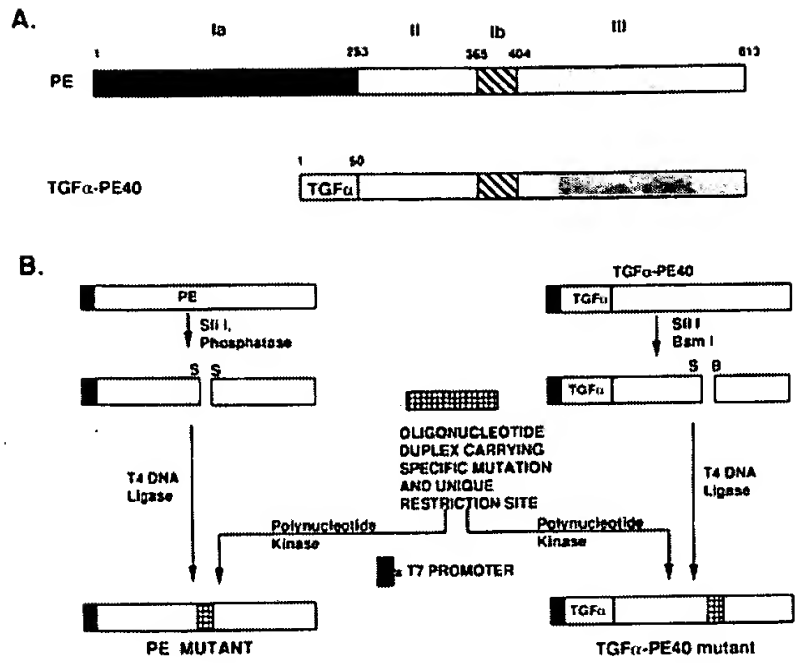


FIGURE 1: (A) Schematic diagram of PE and TGFα-PE40. (B) Scheme for construction of mutant plasmids encoding either PE or TGFα-PE40. S, *Sfi*I; B, *Bsm*I.

of PE. We have found that all of the F helix (aa 361–364) and part of the E helix (aa 346–350) can be deleted but residues at positions 339–345 are required.

Materials and Methods

Bacterial Strains and Plasmids. HB101 (Bethesda Research Laboratories, Gaithersburg, MD) was used for propagation of the plasmids. BL21 (λDE3), which carries an inducible T7 RNA polymerase gene on a λ prophage, was used to express the recombinant proteins (Studier & Moffatt, 1986). The plasmids pVC 387 (encoding TGFα-PE40) and pVC 45f(+)T (encoding native PE) were used as templates in all the constructions (Siegall et al., 1989b; Jinno et al., 1989).

The TGFα-PE40 mutants (pVC387 derivatives) were produced by the insertion of oligonucleotide duplexes into the *Sfi*I–*Bsm*I restriction sites found in the PE gene (Figure 1). The PE mutants [pVC 45f(+) T derivatives] were made by insertion of oligonucleotide duplexes containing the appropriate mutation into the gap formed by a complete *Sfi*I digestion (Figure 1). Several plasmids of each mutant were identified by the presence of a unique restriction site in the new sequence. Additionally, many of the mutants were DNA-sequenced to confirm the presence of the nucleotide substitution.

Expression and Analysis of Recombinant Proteins. Several plasmids from each mutant were expressed into protein by using the inducible T7 promoter system in BL21 (λDE3), and proteins either were used in nonpurified form for certain analysis or were purified from inclusion bodies (TGFα-PE40 derivatives) or from the periplasm fraction (PE and derivatives) (Siegall et al., 1989a,b). Proteins were visualized by electrophoresis on 10% sodium dodecyl sulfate–polyacrylamide gels (SDS–PAGE) using the Laemmli method. Proteins were also tested for authenticity by immunoblot analysis using rabbit antibodies to *Pseudomonas* exotoxin or goat antibodies to TGFα (Biotope, Seattle, WA) and detection using vector stain kits (Vector Laboratories, Burlingame, CA). Proteins were quantitated by using Bradford analysis. ADP-Ribosylation activities of mutant proteins were measured according to Collier and Kandel (1971).

Protein Synthesis Inhibition Analysis. The cytotoxic activities of mutant proteins were compared to that of the native protein using A431 cells (TGFα-PE40) or Swiss-3T3 cells (PE). Protein synthesis was assessed by determining the incorporation of [^{3}H]leucine into cellular protein (Siegall et al., 1989b; Jinno et al., 1989).

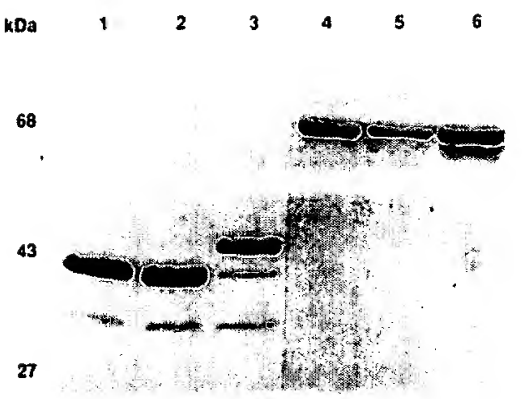


FIGURE 2: Sodium dodecyl sulfate–polyacrylamide gel electrophoresis of TGFα-PE40, PE, and mutant derivatives of both. Samples were applied to 10% SDS–polyacrylamide gels and stained with Coomassie Blue R-250. Lane 1, TGFα-PE40 Δ346–380; lane 2, TGFα-PE40 Δ344–380; lane 3, TGFα-PE40; lane 4, PE; lane 5, PE^{Gln343}; lane 6 PE^{Gln344,345}.

Binding Displacement and Cell-Associated Processing Analysis. ^{125}I-PE binding assays were performed by using Swiss-3T3 cells plated at 1×10^5 cells/mL. Radiolabeled PE (5 ng) was added to cells in the presence of potential competitors at concentrations of 8, 40, and 200 μg as previously described (Jinno et al., 1989). For the processing assay, PE or mutated forms of PE were metabolically radiolabeled with [^{3}H]leucine as described (Ogata et al., 1990). Swiss 3T3 cells (confluent in 150-mm tissue culture dishes) were incubated with labeled toxins at 37 °C for 3 h, washed 5 times with cold PBS, and lysed in 1 mL of RIPA buffer (20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 1% deoxycholic acid, and 0.1% SDS) containing 1 mM PMSF. PE or PE-processed fragments were immunoprecipitated by rabbit anti-PE serum as described (Ogata et al., 1990). Samples were analyzed on 12.5% SDS–PAGE and exposed to X-ray film for 5 days.

Results

Precipitation of TGFα-PE40, PE, and Mutant Proteins. To express TGFα-PE40, PE, and mutant proteins, we transformed *Escherichia coli* BL21 (λDE3) with the appropriate plasmid, cultured the cells, and induced expression of protein with isopropyl β-D-thiogalactoside. TGFα-PE40 and derived proteins were isolated from sucrose-washed spheroplasts and quantitated by comparison of serial dilutions with purified TGFα-PE40 on immunoblots using an antibody to *Pseudomonas* exotoxin. PE and related mutant proteins were expressed from a plasmid containing the OmpA signal sequence and secreted as soluble proteins into the periplasmic space. All proteins were purified by anion-exchange and gel filtration chromatography. Crude and purified proteins were analyzed by SDS gel electrophoresis and by immunoblotting using antibodies to PE (Figure 2).

Cytotoxic Activity of TGFα-PE40 and Related Mutant Proteins. To determine how much of the carboxyl end of domain II was necessary for cell-killing activity, we made deletions of increasing size in domain II of TGFα-PE40. We began making deletion mutants using information from a previous study in which various portions of domains Ib and II were deleted from TGFα-PE40 (Siegall et al., 1989a). We focused on amino acids 337–360 which were demonstrated to contain important residues for cell-killing activity. The various deletion mutants we constructed and a summary of the cytotoxicity of the mutant molecules on A431 cells are shown in Figure 3. Cytotoxicity was assessed by measuring inhibition of protein synthesis. To do this, various dilutions of the mutant

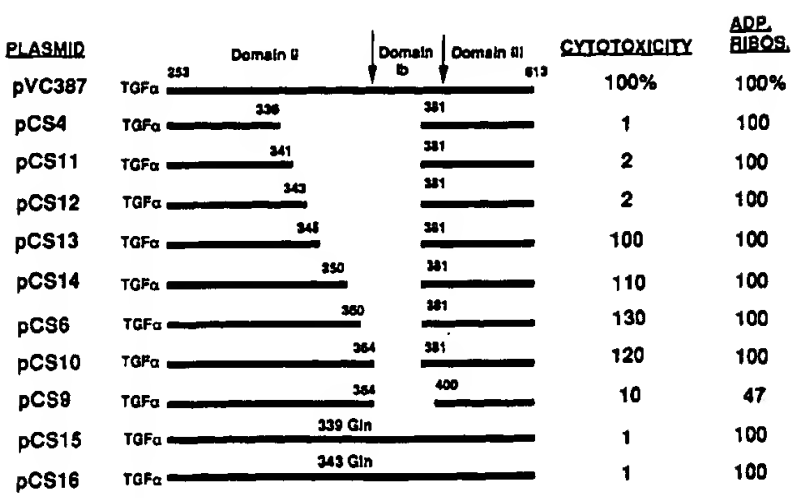


FIGURE 3: Representations of TGFα-PE40 and the various mutant derivatives. Deletion and substitution mutants are shown, and numbers indicate amino acids present or mutated in PE40. The new amino acid in the substitution mutants is indicated next to the residue number. The cDNA for TGFα encodes amino acids 1–50 in all the constructs. Cytotoxicity was determined by comparing the ID_{50} of unmodified TGFα-PE40 to the various mutant forms. ID_{50} is the concentration of toxin which is required to inhibit protein synthesis by 50% as measured by [^{3}H]leucine incorporation after 20-h incubation of A431 cells (1×10^5 cells/mL) with the various toxins. ADP-ribosylation was determined as previously described (Collier & Kandel, 1971).

TGFα-toxins were added to A431 cells and cultured for 20 h followed by a 1-h pulse with [^{3}H]leucine as detailed under Materials and Methods. In each case, we compared the activity of the mutants with unmodified TGFα-PE40. We also measured the ADP-ribosylation activity of each mutant protein (Collier & Kandel, 1971). All the TGFα-toxin mutants had ADP-ribosylation levels identical with that of TGFα-PE40 except for pCS9, described in Hwang et al. (1987) and Figure 3. The results indicate that deletion of amino acids (aa) 346–380 did not reduce the cytotoxicity of TGFα-PE40. However, deletion of two additional residues, aa 344–380, resulted in a molecule which retained only 2% of the cytotoxicity of TGFα-PE40. Deletion mutants extending to amino acids 342 and 337 had almost no cytotoxic activity against A431 cells (Figure 3).

Cytotoxicity of TGFα-PE40 and PE Point Mutants. To define further the important residues in this portion of the toxin molecule, we constructed two point mutations in TGFα-PE40 (Figure 3). The results show that mutations at amino acids 339 and 343, which change alanine to glutamine, eliminated cell-killing activity (Figure 3).

Since we wanted to perform additional analysis in this region of the molecule and also study the intracellular processing of mutant proteins, we made similar mutations in native PE (Figure 4). By generating mutations in native PE, we were able to study the cell-associated processing of the molecules as recently described by Ogata et al. (1990). The processing of chimeric toxins such as TGFα-PE40 has not yet been studied. As with TGFα-PE40, the mutations at amino acids 339 and 343, which changed alanine to glutamine, inactivated the cytotoxic activity of native PE (Figure 4). However, the toxicity of PE was not changed when Ala-343 was changed to leucine or valine (data not shown). Additional point mutations were made in this cluster of amino acids. Surprisingly, when we changed the alanines at amino acids 344 and 345 to glutamine, singularly or together, PE retained full toxicity. Also, when amino acid 341 was mutated from threonine to phenylalanine, the toxin retained activity. The cytotoxicity of the PE point mutations on Swiss 3T3 cells is summarized in Figure 4, and representative dose response curves are shown in Figure 5.

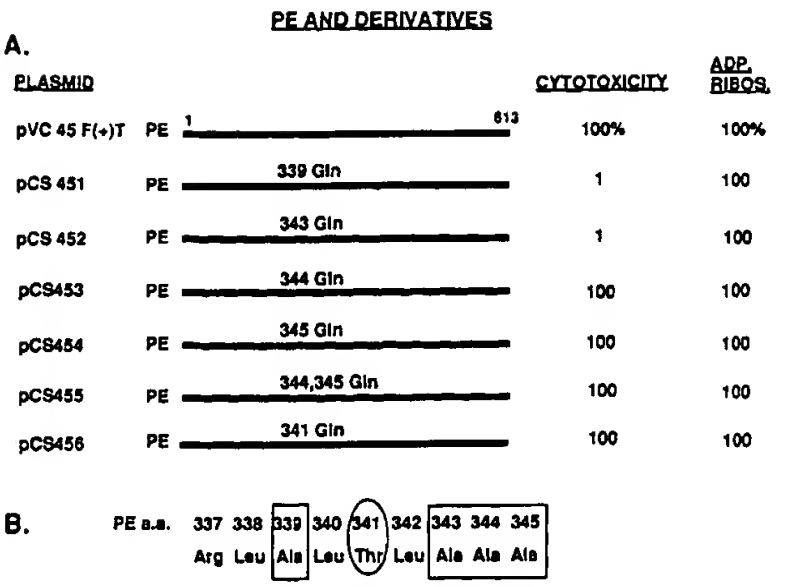


FIGURE 4: (A) Representations of PE and the various mutant forms. Substitution mutants are shown with the new amino acid (Gln in all cases) listed next to the position which was mutated. Cytotoxicity and ADP-ribosylation were done as in Figure 3. (B) The amino acid sequence of PE between residues 338 and 345. The boxes indicated alanine residues which were mutated to glutamine. The circled amino acid indicated a threonine which was mutated.

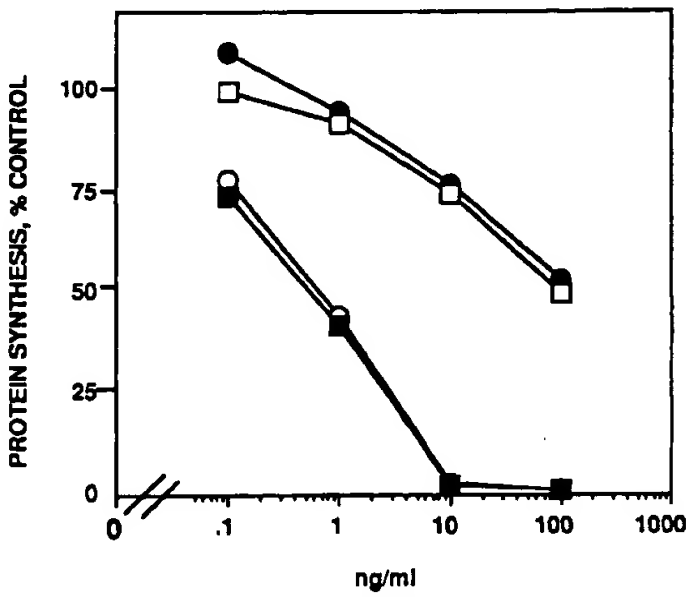


FIGURE 5: Toxic activity of PE and various mutant proteins on Swiss 3T3 cells. PE and derivatives were prepared from the periplasm of *E. coli*. The concentration of the various mutant toxins was estimated by Bradford assay. [^{3}H]Leucine incorporation was measured, and the results are expressed as percent of control cells not treated with PE. PE (○), PE^{Gln343} (●), PE^{Gln339} (□), $PE^{Gln344,345}$ (■).

Binding of PE Mutants. The decrease in cytotoxicity of the PE mutants could be a result of diminished binding to the PE receptor through some indirect affect of domain II on domain Ia. To determine whether the PE mutants had the same affinity as PE for the PE receptor, we compared their displacement of ^{125}I-PE to Swiss 3T3 cells (Figure 6). We added various amounts of competitor (8–200 μg/mL) to 200 ng/mL ^{125}I-PE and incubated for 1 h at 37 °C followed by washing the cells with phosphate-buffered saline and harvesting of the cells for counting. We compared PE and two mutants (PE^{Gln343}; $PE^{Gln344,345}$) which were representative of the mutants generated. A similar degree of displacement of ^{125}I-PE was obtained with PE, PE^{Gln343} and $PE^{Gln344,345}$. Thus, PE and two mutant forms of PE, one, PE^{Gln343}, which inactivates the toxin and a second, $PE^{Gln344,345}$, which does not alter PE cytotoxicity, bind in an identical manner to PE.

Detection of Cell-Associated Processing. PE is proteolytically cleaved after endocytosis, and a 37-kDa C-terminal fragment that contains the ADP-ribosylation activity is translocated into the cytosol (Ogata et al., 1990). To test

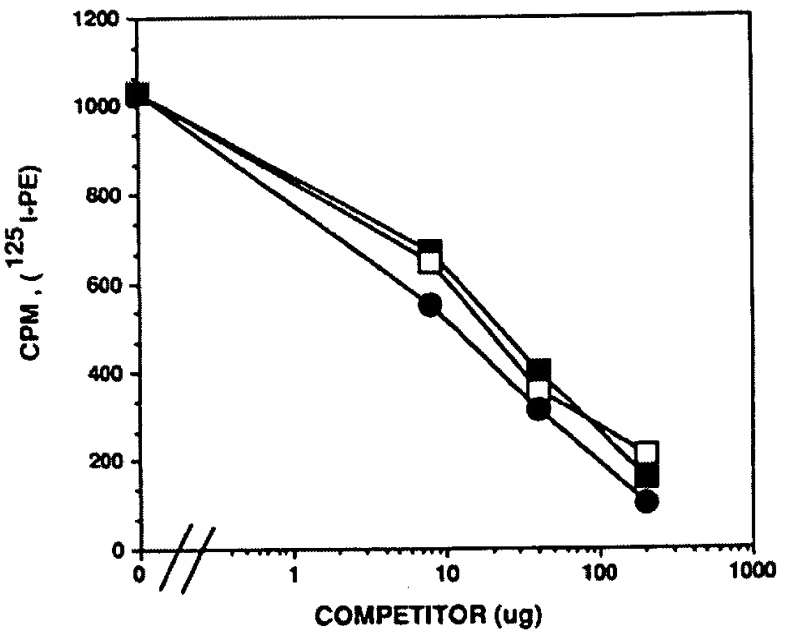


FIGURE 6: PE and PE mutant displacement assay. Binding of PE and two mutant forms was assessed by displacement of trace amounts of ^{125}I-PE. ^{125}I-PE bound to Swiss 3T3 cells was measured as counts per minute. PE (■), $PE^{Gln344,355}$ (□), PE^{Gln343} (●).

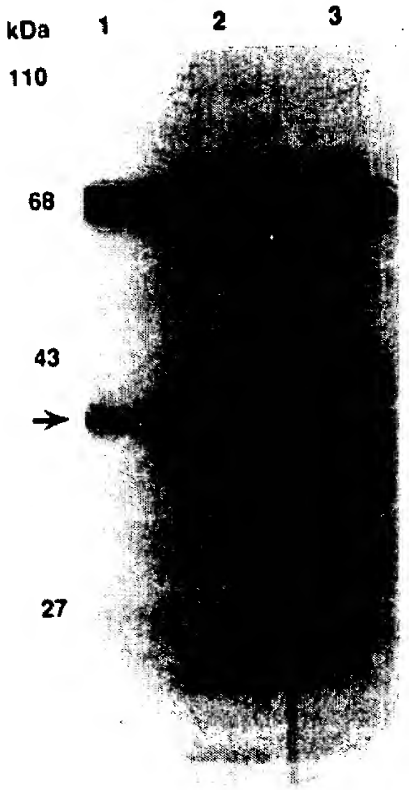


FIGURE 7: Cell-associated processing of radioactive PE and mutants in Swiss 3T3 cells. PE and derivatives were metabolically labeled with [^{3}H]leucine. The radiolabeled proteins were incubated with Swiss 3T3 cells for 3 h, lysed, and immunoprecipitated by using rabbit anti-PE antibodies. Lane 1, PE; lane 2, PE (Gln-343); lane 3, PE (Gly-276). SDS–PAGE was done by using 12.5% gels and reducing conditions. Previously, we showed that PE is first cleaved to produce two fragments, one of 28 kDa, derived from the N-terminus, and one of 37 kDa, derived from the C-terminus. The arrow shows the position of the 37-kDa fragment (which, if the toxin is active, is translocated to the cytosol). The 28-kDa fragment, derived from domain Ia, is further degraded to produce fragments of 25 and 18 kDa (Ogata et al., 1990).

whether the mutant forms of PE which are not cytotoxic to Swiss 3T3 cells are processed in a similar fashion to native PE, we examined the processing of several different PE molecules by Swiss 3T3 cells. We used the PE^{Gln343} mutant and compared its processing to that of PE and PE^{Gly276}, a nontoxic mutant that is not processed appropriately (Figure 7). In this experiment, PE or mutant forms of PE were metabolically radiolabeled with [^{3}H]leucine, purified, and incubated with Swiss 3T3 cells at 37 °C for 3 h. Then the cells were washed, lysed, and immunoprecipitated by rabbit

Table I: Bioactivities of PE and Mutant Molecules

	binding	processing	ADP-ribosylation	cytotoxicity
PE	+	+	+	+
PE^{Gln343}	+	+	+	−
PE^{Gly276}	+	−	+	−
$PE^{Gln344,345}$	+	ND[a]	+	+

[a] ND, not done.

anti-PE serum. The results (Figure 7) indicate that PE and PE^{Gln343} are both processed to a 37-kDa fragment while the mutant PE^{Gly276} is not (Ogata et al., 1990). Therefore, the cytotoxic-deficient mutant PE^{Gln343} can be processed to a 37-kDa fragment. In the PE^{Gln343} mutant, an additional low molecular weight fragment was observed which was derived from domain Ia (Ogata et al., 1990).

Discussion

In this study, we have defined by deletion analysis how much of the carboxyl end of PE domain II is necessary for cytotoxic activity. Previously, we showed that part of domain Ib, amino acids 365–380, could be deleted without loss of activity. We now show that removal of up to 19 aa (Δ346–364) did not result in loss of activity. However, the removal of two additional aa (Δ344–364) did result in a dramatic loss of activity. The α-carbon backbone structure of PE is shown in Figure 8A, and that of domain II is shown in Figure 8B. These deletions remove all of the F helix (aa 360–363) and part of the E helix (aa 333–350). These amino acids, which lie grouped together on the surface of the PE molecule, appear to have no particular role in cytotoxicity. However, they could have a function not detected by our assays.

To gain further information about the E helix, we made point mutations (substitutions) in the essential region that could not be deleted at positions 339, 341, 343, 344, and 345 (Figure 8B). This is a hydrophobic region rich in alanine residues. The mutations changed this small hydrophobic residue to a bulky hydrophilic residue. Two of these mutants (Ala-339 and Ala-343 to Gln) inactivated PE cytotoxicity, while two others (Ala-344 and -345 to Gln) did not inactivate PE (Figures 3 and 4). To determine if these mutations affected other PE functions, we measured ADP-ribosylation and cell binding and found that there were no changes in the mutant molecules as compared to PE (Figures 3–5).

Domain II is important for translocation in a process in which the molecule is first cleaved by a cell-associated protease near residue 279, generating a 37-kDa fragment which is then translocated to the cytosol. Therefore, we examined the processing of the mutant PE^{Gln343} which had lost activity. We found that the mutant molecule was processed to a 37-kDa fragment as was native PE and also retained normal ADP-ribosylating activity (Figure 4). We interpret this result to mean that the mutant molecule (Ala → Gln) generated the same 37-kDa fragment (aa 280–613), as did native PE, but the fragment did not reach the cytosol so that protein synthesis was not arrested. We assume the same result would be obtained with the other inactivating mutations but have not yet examined them. This class of mutants differs from others in domain II such as the change of arginine to glycine at position 276 or 279 which prevents the molecule from being processed normally (Ogata et al., 1990). Currently, their are only a limited number of assays for toxin function, and the results obtained for several mutant proteins are summarized in Table I. Unfortunately, a direct assay for toxin translocation does not yet exist. However, if and when such an assay is developed, we can then examine the role of key amino acids in the E helix.

Table II: Sequence Comparison between Clathrin Light Chain and PE[a]

clathrin	30	31	32	33	34	35	36	37	38	39
light chain	Ala	Ala	Phe	Leu	Ala	Gln	Gln	Glu	Ser	Glu
Pseudomonas	339	340	341	342	343	344	345	346	347	348
exotoxin A	Ala	Leu	Thr	Leu	Ala	Ala	Ala	Glu	Ser	Glu
	↓		↓		↓	↓	↓			
	Gln		Phe		Gln	Gln	Gln			
PE cytotoxicity	-		+		-	+	+			

[a] Amino acids with identity are underlined. Mutation in PE and the cytotoxic activity of the mutant protein are indicated.

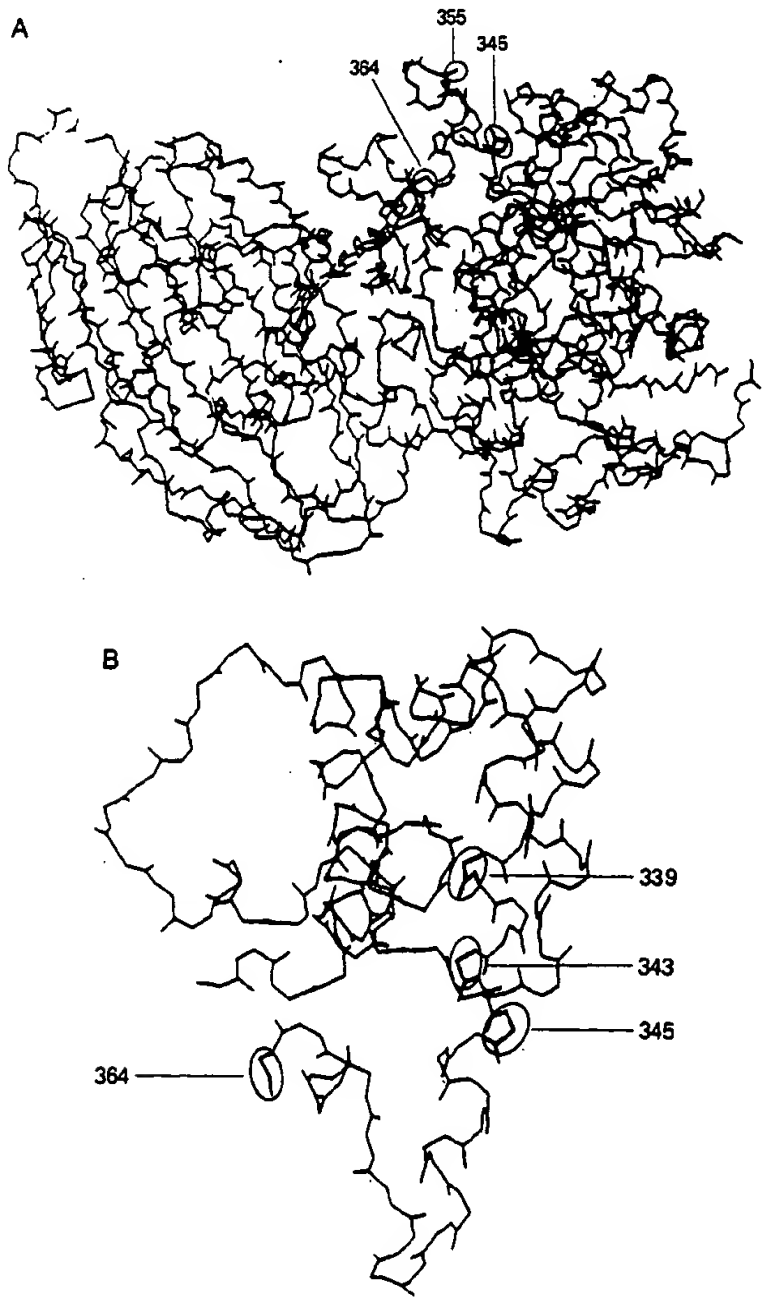


FIGURE 8: Structural representations of *Pseudomonas* exotoxin showing regions where mutations and deletions were made. (A) The entire PE molecule is depicted, and the regions of domains II and Ib which could be deleted are shown (aa 346–380). (B) Structure of PE domain II showing the region from residue 346 to residue 364 which could be removed without affecting cytotoxicity. Amino acids 339 and 343 (both alanines) which are essential for PE cytotoxicity are indicated by their residue number. These two amino acids are on the inside of the helix facing the interior of the PE molecule. Perhaps the inactivating mutations in which glutamine is substituted for alanine at positions 339 and 343 disrupt its structure due to its side chains. Alternatively, the alanines at residues 344 and 345 which can be mutated to glutamine are not on the interior of the E helix. The structures were made by using a Silicon Graphics Iris computer and a GEMM software program developed by Dr. B. K. Lee, NIH.

The region of PE examined in this study, aa 339–345, is rich in hydrophobic residues (Ala, Leu, Thr, Leu, Ala, Ala, Ala). This prompted us to examine whether domain II of PE is related to other proteins that contain hydrophobic stretches of aa residues. The best fit of any protein examined was to clathrin light chain, which had 28.6% identify in a 77 amino acid overlap. When one focuses only on the region of interest in domain II (amino acids 339–348), there is a striking resemblance between the two proteins (Table II). Analysis of the mutations produced (Figures 3 and 4) demonstrates that when we altered an amino acid that was conserved between PE and clathrin light chain, the toxicity of PE was lost. Alternatively, when we mutated an amino acid of PE to an amino acid residue that is found at the corresponding site in clathrin light chain, cytotoxicity was retained (Table II).

It is intriguing that clathrin light chain has significant homology to domain II of PE, since clathrin is found on the cytoplasmic face of coated pits. Clathrin is thought to be involved in facilitating the internalization of cell-surface receptors and to facilitate receptor transport in a process that targets proteins to intracellular compartments (Brodsky, 1988). The region of PE that we have mutated aligns with a region of clathrin light chain that is conserved between all species tested (Jackson & Parham, 1988). The function of this region (aa 23–44 in clathrin light chain) is not known. The central region of clathrin light chain (residues 93–157) is thought to be involved in clathrin heavy-chain binding (Brodsky, 1987). Clathrin light chains also interact with proteins found in the cytoplasm and in vesicles (Brodsky, 1988). It is possible that PE interacts with the same element (perhaps a protein) as clathrin light chains during the translocation process. It is also possible that, coincidentally, this portion of the clathrin light chain has a helical structure and the relationship to PE is more one of common structure than common function.

ACKNOWLEDGMENTS

We thank Drs. Sankar Adhya, R. Heinrickson, and F. Kezdy for their helpful discussions and insights, B. K. Lee for supplying us with the GEMM program, Charlotte Fryling for technical assistance, Elizabeth Lovelace and Annie Harris for cell culture, Dr. Stephen Epstein for critical reading of the manuscript, and Jennie Evans and Althea Gaddis for expert secretarial assistance.

REFERENCES

Allured, V., Collier, R. J., Carroll, S. F., & McKay, D. B. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1320–1324.

Batra, J. K., Jinno, Y., Chaudhary, V. K., Kondo, T., Willingham, M. C., FitzGerald, D. J., & Pastan, I. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8545–8549.

Brandhuber, B. J., Allured, V. S., Falbel, T. G., & McKay, D. B. (1988) *Proteins: Struct., Funct., Genet. 3*, 146–154.

Brodsky, F. M. (1988) *Science 242*, 1396–1402.

Brodsky, F. M., Galloway, C. J., Blank, G. S., Jackson, A. P., Seow, H.-F., Drickamer, K., & Parham, P. (1987) *Nature 326*, 203–205.

Carroll, S. F., & Collier, R. J. (1987) *J. Biol. Chem. 262*, 8707–8711.

Chaudhary, V. K., FitzGerald, D. J. P., Adhya, S., & Pastan, I. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4538–4542.

Chaudhary, V. K., Mizukami, T., Fuerst, T. R., FitzGerald, D. J., Moss, B., Pastan, I., & Berger, E. A. (1988) *Nature (London) 335*, 369–372.

Chaudhary, V. K., Jinno, Y., FitzGerald, D., & Pastan, I. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 308–312.
Collier, R. J., & Kandel, J. (1971) *J. Biol. Chem. 246*, 1496–1503.
Edwards, G. M., Defeo-Jones, D., Tai, J. Y., Vuscolo, G. A., Patrick, D. R., Heimbrook, D. C., & Oliff, A. (1989) *Mol. Cell. Biol. 9*, 2860–2867.
Estworthy, R. S., & Neville, D. M., Jr. (1984) *J. Biol. Chem. 259*, 11496–11504.
FitzGerald, D., & Pastan, I. (1989) *J. Natl. Cancer Inst. 81*, 1455–1463.
Gray, G. L., Smith, D. H., Baldridge, J. S., Harkins, R. N., Vasil, M. L., Chen, E. Y., & Heynecker, H. L. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 2645–2649.
Greenfield, L., Johnson, V., & Youle, R. (1987) *Science 238*, 536–539.
Hwang, J., FitzGerald, D. J. P., Adhya, S., & Pastan, I. (1987) *Cell 48*, 129–136.
Jackson, A. P., & Parham, P. (1988) *J. Biol. Chem. 263*, 16688–16695.
Jinno, Y., Chaudhary, V. K., Kondo, T., Adhya, S., FitzGerald, D. J., & Pastan, I. (1988) *J. Biol. Chem. 263*, 13202–13207.
Jinno, Y., Ogata, M., Chaudhary, V. K., Willingham, M. C., Adhya, S., FitzGerald, D. J., & Pastan, I. (1989) *J. Biol. Chem. 264*, 15953–15959.
Johnson, V. G., & Youle, R. J. (1989) *J. Biol. Chem. 264*, 17739–17744.
Lorberboum-Galski, H., FitzGerald, D., Chaudhary, V., Adhya, S., & Pastan, I. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1922–1926.
Lukas, M., & Collier, R. J. (1988) *Biochemistry 27*, 7629–7632.
Ogata, M., Chaudhary, V. K., FitzGerald, D. J., & Pastan, I. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4215–4219.
Ogata, M., Chaudhary, V. K., Pastan, I., & FitzGerald, D. J. (1990) *J. Biol. Chem. 265*, 20678–20685.
Olsnes, S., & Sandvig, K. (1988) in *Immunotoxins* (Frankel, A. E., Ed.) pp 39–73, Kluwer Academic Publishers, Norwell, MA.
Pastan, I., & FitzGerald, D. (1989) *J. Biol. Chem. 264*, 15157–15160.
Pastan, I., Willingham, M. C., & FitzGerald, D. J. (1986) *Cell 47*, 641–648.
Siegall, C. B., Chaudhary, V. K., FitzGerald, D. J., & Pastan, I. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 9738–9742.
Siegall, C. B., Chaudhary, V. K., FitzGerald, D. J., & Pastan, I. (1989a) *J. Biol. Chem. 264*, 14256–14261.
Siegall, C. B., Xu, Y., Chaudhary, V. K., Adhya, S., FitzGerald, D., & Pastan, I. (1989b) *FASEB J. 3*, 2647–2652.
Studier, F. W., & Moffatt, B. A. (1986) *J. Mol. Biol. 189*, 113–130.
Vitetta, E. S., Fulton, R. J., May, R. D., Till, M., & Uhr, J. W. (1987) *Science 238*, 1098–1104.
Williams, D. P., Snider, C. E., Strom, T. B., & Murphy, J. R. (1990) *J. Biol. Chem. 265*, 11885–11889.
Wozniak, D. T., Hus, L.-Y., & Galloway, D. R. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8880–8884.

Average Membrane Penetration Depth of Tryptophan Residues of the Nicotinic Acetylcholine Receptor by the Parallax Method†

Amitabha Chattopadhyay‡ and Mark G. McNamee*
Department of Biochemistry and Biophysics, University of California, Davis, California 95616
Received March 26, 1991

ABSTRACT: The membrane penetration depths of tryptophan residues in the nicotinic acetylcholine receptor from *Torpedo californica* have been analyzed in reconstituted membranes containing purified receptor and defined lipids. Dioleoylphosphatidylcholine and three spin-labeled phosphatidylcholines with the nitroxide group at three different positions on the fatty acyl chain were used for reconstitution of the receptor. The spin-labeled phospholipids serve as quenchers of tryptophan fluorescence. Differential quenching of the intrinsic fluorescence of the acetylcholine receptor by the spin-labeled phospholipids has been utilized to analyze the average membrane penetration depth of tryptophans by the parallax method [Chattopadhyay, A., & London, E. (1987) *Biochemistry 26*, 39–45]. Analyses of the quenching data indicate that the tryptophan residues on the average are at a shallow location (10.1 Å from the center of the bilayer) in the membrane. In addition, the generally low levels of quenching imply that the majority of tryptophan residues are located in the putative extramembranous region of the receptor. These results are consistent with several proposed models for the tertiary structure of the acetylcholine receptor and are relevant to ongoing analyses of the overall conformation and orientation of the acetylcholine receptor in the membrane.

Due to the inherent difficulty of crystallizing membrane proteins, most analyses of membrane protein structure have utilized indirect biophysical techniques with an emphasis on spectroscopic methods. One such analysis involves determination of membrane penetration depth, the location of a molecule or a specific site within a molecule in relation to the membrane surface. Knowledge of the precise depth of a membrane-embedded group or molecule helps define the conformation and topology of membrane proteins and probes. Fluorescence has been one of the most widely used techniques to determine depth. Both long-range dipole–dipole (Forster)

†This work was supported by Grant NS13050 from the National Institute of Neurological Disorders and Stroke. A preliminary version of some of the results in this paper was presented at the 33rd Annual Meeting of the Biophysical Society, Cincinnati, OH, Feb 12–16, 1989.
*Address correspondence to this author.
‡Present address: Centre for Cellular and Molecular Biology, Uppal Rd., Hyderabad 500 007, India.

0006-2960/91/0430-7159$02.50/0

Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 2077–2081, March 1990
Biochemistry

GAL4 transcription factor is not a "zinc finger" but forms a $Zn(II)_2Cys_6$ binuclear cluster

(¹H NMR/metalloproteins)

TAO PAN AND JOSEPH E. COLEMAN

The Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06510

Communicated by Bert L. Vallee, December 26, 1989

ABSTRACT **The DNA-binding domain of the transcription factor GAL4, consisting of the 62 N-terminal residues and denoted GAL4(62*), contains a Cys-Xaa_2-Cys-Xaa_6-Cys-Xaa_6-Cys-Xaa_2-Cys-Xaa_6-Cys motif, which has been shown previously to bind two Zn(II) or Cd(II) ions. Binding of Zn(II) or Cd(II) is essential for the recognition by GAL4 of the specific palindromic DNA sequence to which it binds upstream of genes for galactose-metabolizing enzymes, the UAS_G sequence. On the basis of the ^{113}Cd NMR chemical shifts of the two bound $^{113}Cd(II)$ ions, we propose a binuclear cluster model for this Zn(II)-binding subdomain. ^1H–^{113}Cd heteronuclear multiple-quantum NMR spectroscopy and phase-sensitive double-quantum filtered ^1H correlation spectroscopy of the $^{112}Cd(II)$- and $^{113}Cd(II)$-substituted GAL4(62*) derivatives provide direct evidence that the two bound $^{113}Cd(II)$ ions are coordinated only by the six cysteine residues, two of which form bridging ligands between the two $^{113}Cd(II)$ ions. The latter can be identified from the pattern of ^1H–^{113}Cd *J* coupling. Thus a binuclear metal ion cluster rather than a "zinc finger" is formed by the six cysteine residues of the GAL4 DNA-binding domain. This model can be directly applied to eight other fungal transcription factors which have been shown to contain similarly spaced Cys_6 clusters. ^1H NMR spectra of apo-GAL4(62*) suggest conformational fluctuation of the metal-binding subdomain upon removal of Zn(II) or Cd(II). Both $Cd(II)_2$- and $Zn(II)_2$-containing species of GAL4 can be formed, and the similar ^1H NMR spectra suggest similar conformations.**

The GAL4 transcription factor from *Saccharomyces cerevisiae* is a protein of 881 amino acids required for the transcriptional activation, in the presence of galactose, of the genes coding for galactose-metabolizing enzymes (1). The DNA-binding domain, which recognizes a specific nucleotide sequence (UAS_G) upstream of the promoters for these genes, has been localized to the N-terminal 62 amino acids (2–5). This domain of GAL4 contains a cysteine-rich amino acid sequence, Cys^{11}-Xaa_2-Cys^{14}-Xaa_6-Cys^{21}-Xaa_6-Cys^{28}-Xaa_2-Cys^{31}-Xaa_6-Cys^{38}, conserved among a group of nine transcription factors isolated from fungi (6). The cysteine cluster in GAL4 and related transcription factors has been proposed to form a single "zinc finger" containing a tetrahedral Zn(II) coordination complex (7), similar to that described for transcription factor TFIIIA from *Xenopus* oocytes (8). The GAL4 "finger" would have 4 cysteine residues as ligands rather than the 2 cysteine and 2 histidine residues found in TFIIIA and related proteins (7, 8).

We have shown previously that cloned N-terminal fragments of GAL4 consisting of 149 and 62 amino acids, denoted GAL4(149*) and GAL4(62*), contain Zn(II) (4, 5). The Zn(II) is essential for the binding of these fragments to the UAS_G DNA sequence, and Zn(II) can be replaced by Cd(II) without loss of specific DNA binding (4, 5). A single zinc-finger model of GAL4 predicts that the protein contains 1 Zn(II) ion per molecule, but both GAL4(149*) and GAL4(62*) can bind 2 Zn(II) or Cd(II) ions per molecule (4, 5). $^{113}Cd(II)$ substitution of Zn(II)-binding sites in metalloproteins has served as a powerful probe of the structure of these Zn(II) sites, primarily because of the great sensitivity of ^{113}Cd NMR signals to the nature of the ligands to the metal ion. ^{113}Cd NMR in combination with two-dimensional (2D) ^1H NMR has led to the determination of the complete solution structure of the small Zn(II)- and Cd(II)-binding protein metallothionein (9, 10). This protein is the only example of a Zn(II) or Cd(II) cluster complex found in nature in which several of the cysteine residues act as bridging ligands between two metal ions (10). Metallothionein contains 7 metal ions located in two clusters (11).

Our initial ^{113}Cd NMR study of Cd_2GAL4 confirmed the presence of two Cd(II)-binding sites in both GAL4(149*) and GAL4(62*) which induce ^{113}Cd chemical shifts of 707 and 669 ppm, consistent with ligation of each $^{113}Cd(II)$ to at least 3 and probably 4 sulfur atoms (4, 5). We have now used 2D ^1H NMR techniques to demonstrate that the only ligands to the 2 $^{113}Cd(II)$ ions in $^{113}Cd_2GAL4(62*)$ are the 6 cysteine residues. The ^1H–^{113}Cd *J* coupling patterns of the cysteine β protons are best explained by a binuclear cluster in which two of the cysteine residues are bridging ligands between the two metal ions. The NMR data supporting this model of the DNA-binding domain of GAL4 are presented in this paper.

MATERIALS AND METHODS

Cloning, Overproduction, and Purification of GAL4(62*). The GAL4(62*) clone was obtained by insertion of a *Spe* I stop-codon linker, 5′-CCCGGCTAGACTAGTCTAGC-CGGG, into the *Hin*fI site of codon 59 in pTPT7G1 (4). The *Spe* I linker replaces the natural sequence Leu-Glu at residues 61–62 by Leu-Asp. We have termed our construct GAL4(62*). Overproduction and purification of GAL4(62*) was as described for GAL4(149*) (4) except that GAL4(62*) was eluted at standard column buffer plus 250 mM NaCl from a Bluegel agarose column (Bio-Rad).

Preparation of Apo- and Cd(II)GAL4(62*). Purified GAL4(62*) contains 1–2 atoms of Zn(II) per molecule. Apo-GAL4(62*) can be obtained as described for apo-GAL4(149*) (4). Cd(II)GAL4(62*) can be prepared by addition of Cd(II) to the apoprotein at pH 8.0 in the presence of excess 2-mercaptoethanol. Alternatively, Zn(II) can be exchanged with Cd(II) as follows: A mixture of 3-fold molar excess of Cd(II) over Zn(II)GAL4(62*) in the presence of excess 2-mercaptoethanol is incubated at room temperature (22°C) for 12 hr. Dialysis against metal-free buffer then removes more than 90% of the Zn(II), which has become free. The exchange

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: 2D, two-dimensional; HMQC, heteronuclear multiple-quantum correlated spectroscopy; COSY, correlated spectroscopy; DQF-^1H-COSY, 2D double-quantum filtered proton COSY.

process can be repeated to yield GAL4(62*) with two tightly bound Cd(II) ions.

^{1}H NMR. The NMR spectroscopy was performed on a Bruker (Billerica, MA) AM-500 spectrometer at 35°C. All samples were stored under nitrogen to prevent oxidation of SH groups. $^{112}Cd_2$- and $^{113}Cd_2$GAL4(62*) were obtained as described above, lyophilized, and redissolved in ^2H_2O. The concentrations for both species were about 3 mM. Apo- and Zn(II)GAL4(62*) were exchanged into buffer in ^2H_2O (50 mM sodium phosphate/100 mM NaCl, pH 7.8) by a Sephadex G-25 size-exclusion spun column. The acquisition parameters are listed in the figure captions. Chemical shifts are plotted relative to sodium trimethylsilyltetradeuteriopropionate. The Fourier transform NMR program by Dennis Hare (Hare Research, Woodinville, WA) was used for processing the 2D NMR spectra.

RESULTS

^{1}H–^{113}Cd Heteronuclear Multiple-Quantum Correlated Spectroscopy (HMQC) of ^{113}Cd(II)GAL4(62*). To identify the protons of GAL4(62*) coupled to the bound ^{113}Cd(II) (i.e., protons within three bonds of the ^{113}Cd nucleus), ^{1}H–^{113}Cd HMQC was carried out on $^{113}Cd_2$GAL4(62*) (9). The chemical shifts of all protons observed to be coupled to ^{113}Cd(II) are shown in Fig. 1. Their chemical shifts, 3.8–2.6 ppm, are consistent with the assignment of all the resonances to the β protons of the 6 cysteine residues in the protein. No protons from other residues such as those of the CH_3 group of the N-terminal methionine are coupled to the ^{113}Cd(II) ions.

2D ^{1}H Correlated Spectroscopy (COSY) of GAL4(62*). The resonances of individual β protons coupled to the protein-bound ^{113}Cd(II) ions can be separated and identified by comparing the phase-sensitive double-quantum filtered (DQF)-^{1}H-COSY of $^{112}Cd_2$ ($I = 0$) and $^{113}Cd_2$ ($I = ½$) -GAL4-(62*). Resolution of the additional ^{1}H–^{113}Cd scalar coupling of the α–β and β–β cross-peaks of the cysteine side chains allows the precise determination of the number of cysteine residues coordinated to Cd(II), and the magnitude and pattern of the J coupling observed give information on the structure of the ^{113}Cd(II)-binding sites.

^{1}H–^{113}Cd Coupling Patterns for the α–β Cross-Peaks of Cysteine Ligands in $^{113}Cd_2$GAL4(62*). Regions of the phase-sensitive DQF-^{1}H-COSY containing the α-β cross-peaks for 5 of the 6 cysteine residues (designated Cys 2–6) are shown in Fig. 2 *A* and *B* for ^{112}Cd(II)$_2$GAL4(62*) and ^{113}Cd(II)$_2$-GAL4(62*), respectively. In the case of Cys 2 the α–β^a and α–β^b cross-peaks are degenerate due to the similar chemical shifts of the two β protons (Table 1). The change from ^{112}Cd to ^{113}Cd is expected to induce additional scalar coupling only to Cys α–β and β^a–β^b cross-peaks. The α–β cross-peaks of Cys 1 lie outside of the spectral region shown in Fig. 2 and are shown in Fig. 3*A*. Supportive confirmation for the assignment of Cys 4 in Fig. 2 is shown in Fig. 3*B*, taken from the other half of the COSY spectrum. The chemical shifts of the cysteine β^a and β^b protons determined from the phase-sensitive ^{1}H COSY are consistent with those obtained from the ^{1}H–^{113}Cd HMQC (Fig. 1, Table 1). The ^{1}H–^{113}Cd coupling

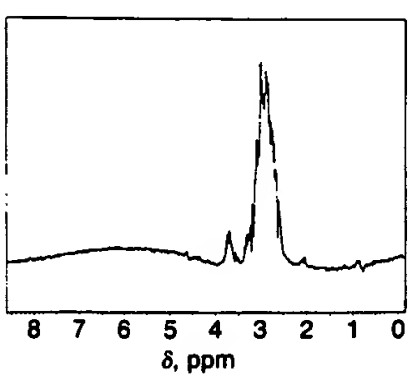


FIG. 1. ^{1}H–^{113}Cd HMQC [$90°_x$(^{1}H)–τ–$90°_\phi$(^{113}Cd)–$t_1/2$–$180°$(^{1}H)–$t_1/2$–$90°_x$(^{113}Cd)–τ–FID] of ^{113}Cd(II)GAL4-(62*) in ^2H_2O; τ = 8 ms. The number of transients was 1024 and the relaxation delay was 1.5 s.

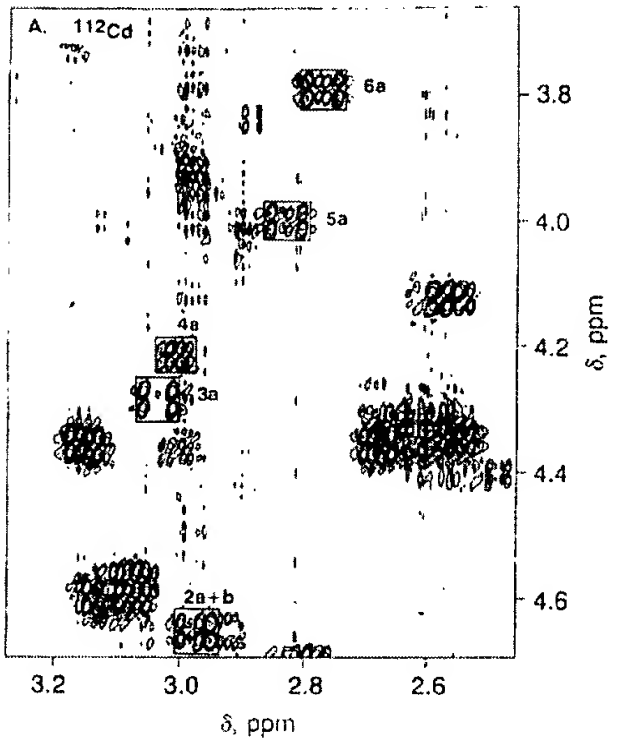


FIG. 2. Comparison of the phase-sensitive DQF-^{1}H-COSY of $^{112}Cd_2$GAL4(62*) (*A*) and $^{113}Cd_2$GAL4(62*) (*B*). For both experiments sweep width was 2790 Hz and 2048 data points in the t_2 and 512 experiments in the t_1 domain were collected. The spectra represent 64 scans each. The relaxation delay was 1.7 s. α–β^a,β^b cross-peaks are shown. Numbers by boxed peaks refer to cysteine residues arbitrarily designated as 1–6. Note that in Figs. 2 and 4 the contours are not phase coded.

observed for the β^a protons of all 6 cysteine residues of GAL4(62*) confirms that all 6 residues are coordinating at least one of the two Cd(II) ions in the protein.

^{1}H–^{113}Cd Coupling Patterns for the β^a–β^b Cross-Peaks for Cysteine Ligands in GAL4(62*). The β^a–β^b cross-peaks of Cys 3, Cys 4, Cys 5, and Cys 6 are located in the section of the $^{112}Cd_2$GAL4(62*) and $^{113}Cd_2$GAL4(62*) ^{1}H COSY shown in Fig. 4 *A* and *B*, respectively. The detailed ^{1}H–^{113}Cd coupling pattern of the β^a–β^b cross-peaks of Cys 5 and Cys 6 are shown in Fig. 5. Metallothionein provides the best guide to the ^{1}H–^{113}Cd coupling expected for the $\alpha\beta$ and $\beta^a\beta^b$ ^{1}H-COSY cross-peaks of cysteine residues coordinated to ^{113}Cd(II) (12). In metallothionein there are 20 cysteine residues coordinated to Cd(II), 8 of which bridge two Cd(II) ions (12). From the ^{1}H

Table 1. Chemical shifts and ^1H–^1H and ^1H–^{113}Cd coupling constants for six cysteine residues of GAL4(62*)

No. of spin system	$^{113}Cd_2$GAL4(62*) at 35°C, pH 8.0										Zn_2GAL4(62*) at 35°C, pH 5.4		
	Chemical shift, ppm			^1H–^1H coupling constants, Hz			^1H–^{113}Cd coupling constants, Hz				Chemical shift, ppm		
	H_α	H_{β^a}	H_{β^b}	$J_{\alpha\beta^a}$	$J_{\alpha\beta^b}$	$J_{\beta^a\beta^b}$	$J_{\beta^aCd^1}$	$J_{\beta^bCd^1}$	$J_{\beta^aCd^2}$	$J_{\beta^bCd^2}$	H_α	H_{β^a}	H_{β^b}
1	4.94	3.73	3.60	11	10	15	25	10			4.67	3.60	3.50
2	4.64	2.95	2.95	—	—	—	≈52	—			4.65	2.84	2.84
3	4.33†	3.04	2.67	11	<2	22	55	26			4.29	3.04	2.55
4	4.22	2.99	3.14	7	<2	16	20	45			4.12	2.88	—
5	3.98	2.82	3.35	10	<2	14	28	14	5‡	3	3.98	2.75	3.25
6	3.75	2.77	3.26	10	<2	14	29	16	2‡	3	3.73	2.73	3.17

Note that by definition β^a is the proton with the larger homonuclear spin–spin coupling (12). A — indicates not obtainable due to lack of resonances or degeneracy of the chemical shifts.
†Chemical shift differs from that of $^{112}Cd_2$GAL4(62*) (H_α = 4.29). The reason for such differences is not clear.
‡From α–β^a cross-peaks.

COSY of $^{113}Cd_7$ metallothionein, it is apparent that in the presence of a bridging cysteine, coupling of the β protons to one of the coordinated ^{113}Cd nuclei is generally of much greater magnitude than to the other. In fact, only two of the shared cysteine residues in metallothionein can be readily identified in the ^1H COSY of the $^{113}Cd_7$ derivative (12). The β^a–β^b cross-peaks of Cys 5 and Cys 6 of GAL4(62*) show splitting patterns similar to the pattern shown by the one "classical" shared cysteine (C14 = Cys^{44} in the sequence) in metallothionein. However, the discernible $^{113}Cd^2$–β^b coupling in ^{113}Cd(II)GAL4(62*), expressed as increased vertical separation of the multiplet on shifting from ^{112}Cd to ^{113}Cd, is quite small, 3–4 Hz for both Cys 5 and Cys 6. Although a $J_{H\beta^a{}^{113}Cd^2}$ is not indicated for either Cys 5 or Cys 6 in Fig. 5 ($J_{H\beta^a{}^{113}Cd^1}$ = 28 and 29 Hz, respectively), the splitting of the α–β^a cross-peaks of these same cysteine residues is ≈33 and 31 Hz (Fig. 2*B*). This difference in coupling of $^{113}Cd^1$ to β^a as measured from the β^a–β^b vs. the α-β^a cross-peak suggests that significant coupling (≈2–5 Hz) of β^a to $^{113}Cd^2$ is present. In addition, partial cancellation of the multiplets in the α–β cross-peaks of Cys 5 and Cys 6 is also suggestive of finite coupling constants to a second ^{113}Cd ion. Such coupling is more directly manifest in the α–β^a than in the β^a–β^b cross-peak. The ^1H–^{113}Cd coupling constants for GAL4(62*) derived from the phase-sensitive ^1H COSY vary from 2 to 55 Hz (Table 1). The range of values is comparable to that observed for the $^{113}Cd_7$ derivative of metallothionein (12). The β^a–β^b cross-peak of Cys 1 is located downfield of the section of the COSY spectrum shown in Fig. 4 and has been used to measure the ^{113}Cd–$^1H_\beta$ coupling constants given in Table 1. There is no β–β cross-peak for Cys 2 due to the degeneracy of the chemical shifts of the β protons (Table 1).

We have shown previously by circular dichroism that removal of Zn(II) causes substantial conformational change in GAL4(63) (5). Considerable chemical shift changes and reduction of linewidth of several signals are observed when Zn(II) or Cd(II) is bound to apo-GAL4(62*) (Fig. 6). The protons of Zn(II)- and Cd(II)GAL4(62*) have very similar chemical shifts, suggestive of nearly identical conformations for both species.

DISCUSSION

The N-terminal fragments of the transcription factor GAL4, either 147 or 63 amino acids in length, form domains which fold independently and bind to the specific DNA sequence, UAS_G, recognized by GAL4 (4, 5). Both fragments bind 1 to 2 Zn(II) ions, and they maintain the binding of the second Zn(II) when the free Zn(II) concentration is 5 μM (5). When the Zn(II) is replaced with ^{113}Cd(II), the ^{113}Cd NMR of both fragments shows two ^{113}Cd NMR signals, each of which integrates to one ^{113}Cd(II) ion (4, 5). Thus there are two metal-ion-binding sites on the DNA-binding domain of GAL4.

GAL4(62*) has a high-resolution proton NMR spectrum uncomplicated by the oligomerization shown by GAL4(149*) when it is present at the concentrations of protein required for ^1H NMR (5). The GAL4(62*) clone produces a protein structurally and functionally identical to GAL4(63) obtained from GAL4(149*) by partial proteolysis (5). GAL4(62*) is soluble to concentrations of several millimolar and thus ideal for 2D ^1H-NMR studies. The fact that cloned GAL4(62*) is highly soluble, while GAL4-(1–74) is not (5), also suggests that GAL4(62*) forms an independent domain. The ^1H–^{113}Cd HMQC establishes that only the six cysteine residues appear to be ligands to metal ions (Fig. 1). The DQF-^1H-COSY

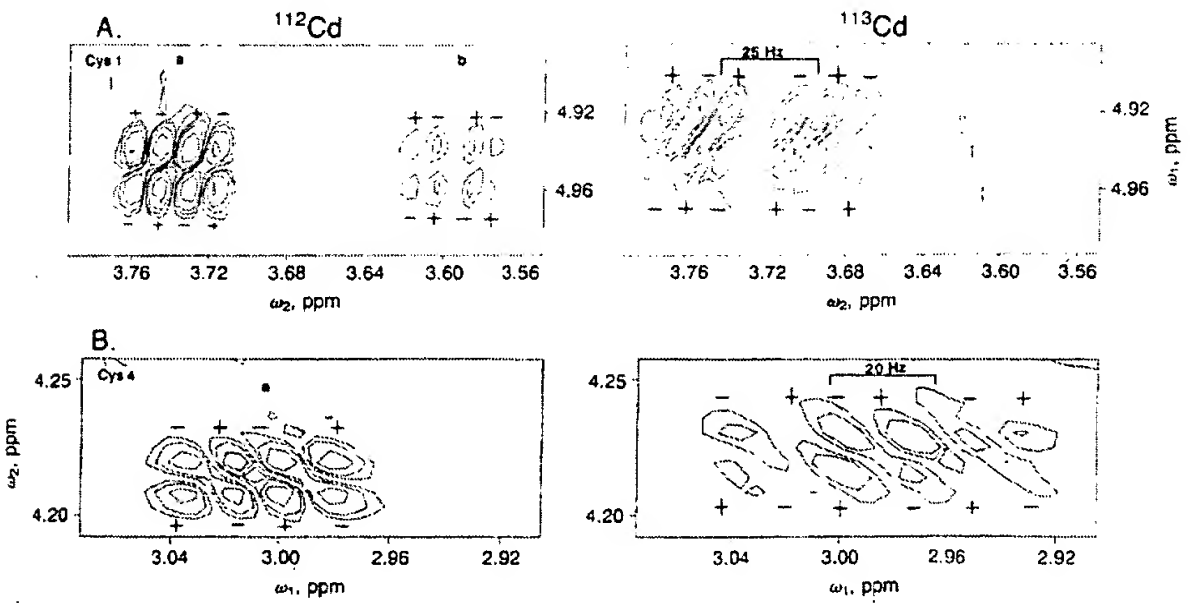


FIG. 3. Phase-sensitive DQF-^1H-COSY regions showing α–β^a,β^b cross-peaks of Cys 1 (*A*) and α–β^a cross-peak of Cys 4 (*B*) (see text). For Cys 1 and ^{113}Cd, additional splitting makes resonance 1b of low intensity.

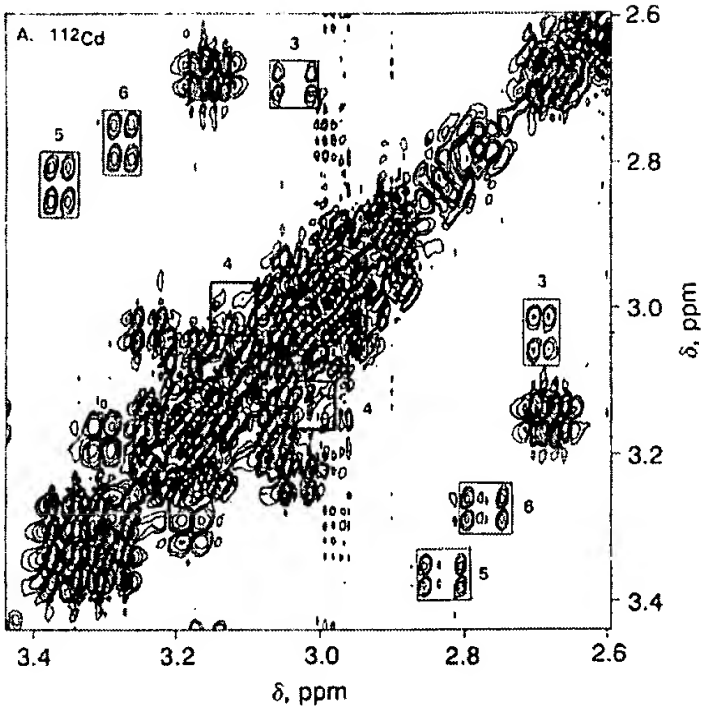


FIG. 4. Phase-sensitive DQF-^{1}H-COSY as in Fig. 2. β^a–β^b cross-peaks of Cys 3, Cys 4, Cys 5, and Cys 6 are shown. (*A*) $^{112}Cd_2$GAL4(62*). (*B*) $^{113}Cd_2$GAL4(62*). The cross-peak for Cys 4 in both derivatives and half the cross-peak of Cys 3 in the ^{113}Cd derivative are badly overlapped with other resonances, but they can be separated in expanded and phase-coded plots. The expanded plots were used to measure the coupling constants given in Table 1.

shows that two of these cysteine residues (designated 5 and 6) appear to be bridging ligands between the two ^{113}Cd(II) ions (Figs. 4 and 5). Among the six ^{113}Cd(II)-coordinated cysteine residues of GAL4(62*), Cys 5 and 6 must occupy almost identical and unique coordination environments, since they both have nearly identical and relatively large coupling of both β protons to one of the ^{113}Cd(II) ions (Fig. 5), a pattern which is also observed for the two well-resolved β^a–β^b cross-peaks of bridging cysteine in metallothionein (12). One can conclude that the torsional angles describing the configuration of the three-bond fragment between the β protons and the coordinated ^{113}Cd(II) ions must be the same for Cys 5 and Cys 6. A possible model of the binuclear complex (cloverleaf) formed by GAL4(62*) is shown in Fig. 7. This model assumes Cys^{21} and Cys^{38} are the bridging ligands, which places them in equivalent configurations, as

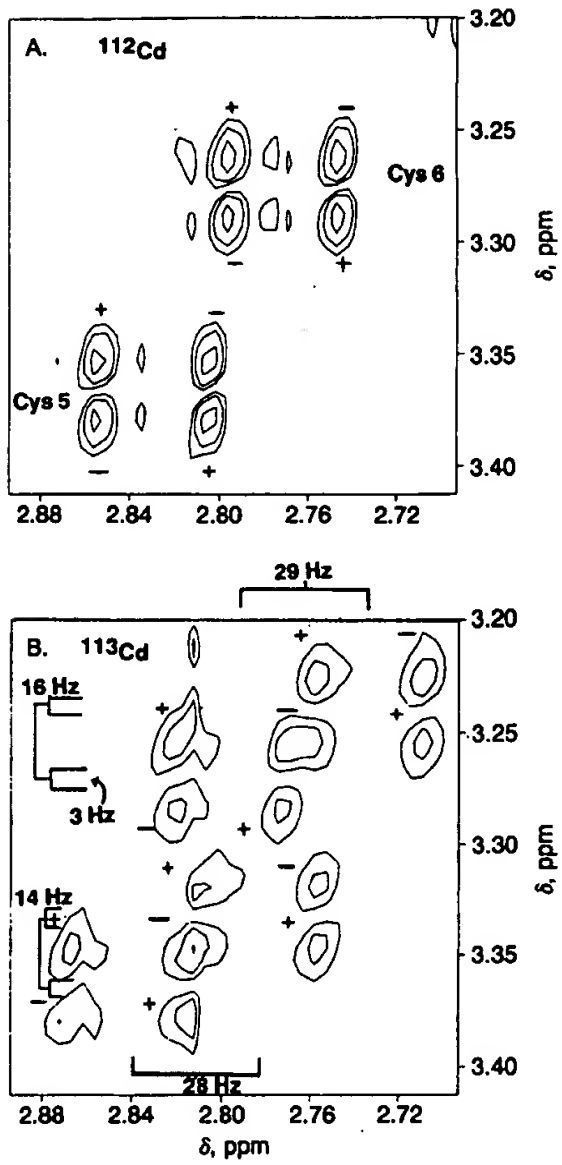


FIG. 5. Expanded region from Fig. 4 showing the coupling patterns derived from the β^a–β^b cross-peaks of Cys 5 and Cys 6.

their *J* coupling suggests. The model building suggests that the five peptide loops necessitated by the cluster structure pack most easily with this choice of bridging ligands. A preliminary sequential assignment of Zn(II)GAL4(62*) supports the assignment of Cys 5 and 6 to residues 38 and 21, respectively, but this result must be considered tentative.

Sequential assignment of the GAL4(62*) spectrum by 2D NMR techniques is not yet complete. While the ^{1}H–^{113}Cd heteronuclear *J* coupling will help establish landmarks in the spectra of this sequence, which has an abundance of lysine and arginine residues, at present its most important aspect is the establishment of the two-metal-ion cluster structure. From the complete COSY and nuclear Overhauser enhancement (NOE) spectroscopy (NOESY) spectra in H_2O, all the NH–H_α connectivities can be identified. Strong NH_i–NH_{i+1} NOEs in the NOESY spectrum of Zn(II)GAL4(62*) have established two separate short stretches of α-helix (unpublished results). Thus a large part of the backbone of GAL4(62*) must be unstructured or β-sheet-like, not incompatible with the model shown in Fig. 7. The regions N-terminal (residues 1–10) and C-terminal (residues 49–62) to the cluster subdomain are likely to form the α-helices suggested to be present by the NOESY data on Zn(II)GAL4(62*) and the earlier circular dichroism studies (5).

The model of GAL4 in Fig. 7 places the preponderance of positive charge, six lysine and one arginine residues, on one side of the cluster formed by the two loops Cys^{14}–Cys^{21} and Cys^{21}–Cys^{28}. Comparison of the sequences shows that the positive charge distribution in the Cys^{14}–Cys^{21} loop is main-

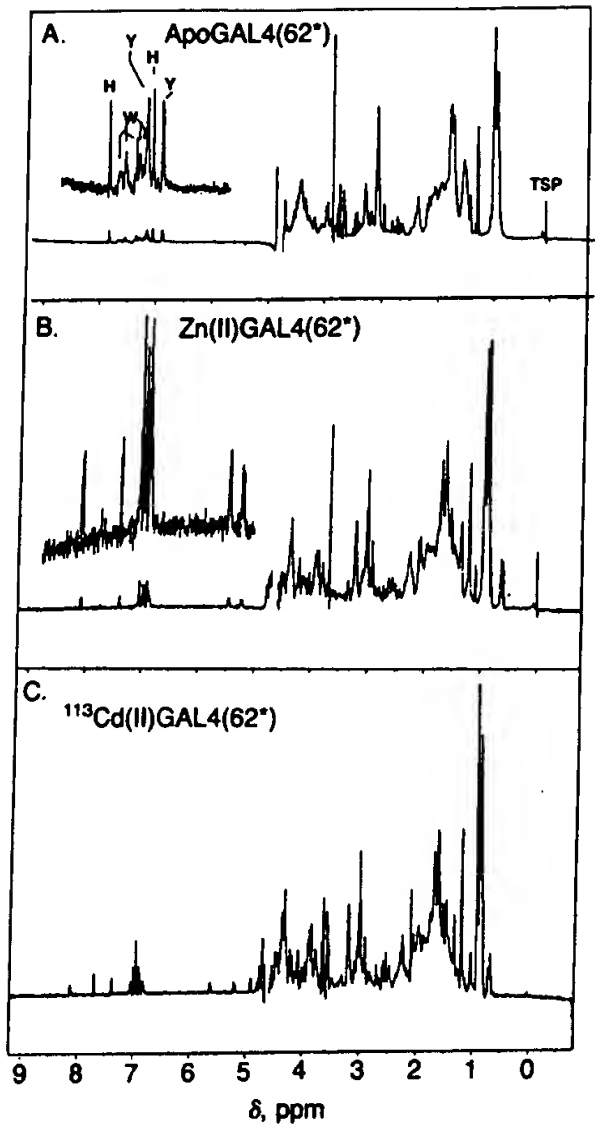


FIG. 6. ^{1}H NMR spectra of a 0.9 mM solution of apo-GAL4(62*) (*A*), Zn(II)GAL4(62*) (*B*), and ^{113}Cd(II)GAL4(62*) (*C*). The number of transients is 16 and the relaxation delay was 2.5 s. TSP, sodium trimethylsilyltetradeuteriopropionate.

tained in all nine fungal transcription factors of this class; in fact, the positions of three of the lysine residues are absolutely conserved in this loop. This suggests that these residues within the cluster (Cys^{11}–Cys^{38}) may function in nonspecific DNA binding and that the recognition determinants may be outside this region. The specific contact region could be provided by a turn and α-helix following the most C-terminal cysteine residue. The specific DNA recognition could involve a cloverleaf–turn–helix structure with turns in the region of Pro^{42} and Pro^{48}. Genetic studies already suggest that amino acid

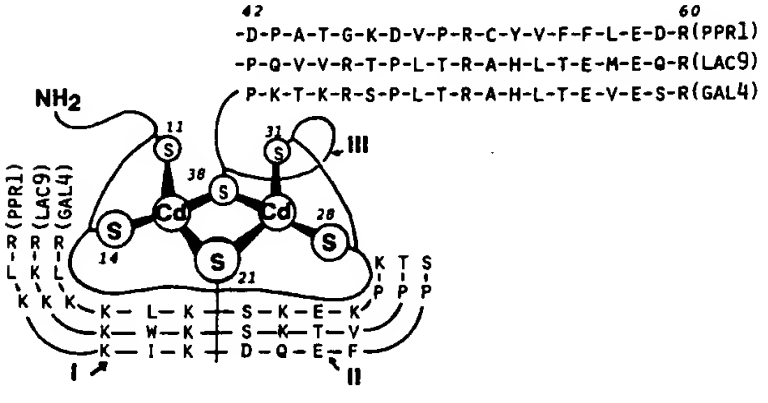


FIG. 7. Model of the $Zn(II)_2Cys_6$ or $Cd(II)_2Cys_6$ cluster (cloverleaf) of the DNA-binding domain of GAL4. Portions of the amino acid sequences of transcription factors LAC9 and PPR1 are also shown. Loops I, II, and III each contain six amino acid residues.

residues within the $Zn(II)_2Cys_6$ cluster of GAL4 are not directly involved in specific DNA recognition (13).

Partial amino acid sequences of two other fungal transcription factors in the regions of the DNA-binding domains corresponding to that of GAL4(62*) are compared in the model (Fig. 7). One of them, LAC9, must be homologous to GAL4, since they both recognize the UAS_G sequence (14, 15). The second, PPR1, does not recognize this sequence but maintains the Cys_6 motif (16). All three factors maintain similar charge and amino acid sequences within the metal ion cluster. As shown for PPR1, the most marked deviation of sequence in this group of transcription factors occurs in the region immediately downstream from the cluster. Of the eight other factors, only the more closely related LAC9 has a sequence in this region similar to that of GAL4 (Fig. 7). This region of the DNA-binding domain may participate in the sequence-specific DNA interactions.

A binuclear zinc cluster complex is a structure different from those previously proposed for the zinc-finger transcription factors. The highly conserved arrangement of cysteine residues in GAL4 and the other fungal transcription factors almost certainly means that this structure is also present in the others. Whether such a binuclear complex or a variant of it is present in other so-called Cys_2Cys_2 zinc-finger transcription factors, such as the steroid receptor proteins (containing a total of nine conserved cysteine residues), remains speculative. Cd(II) forms a stable binuclear cluster with GAL4 and the binding is highly cooperative (4, 5). Not surprisingly, the Zn(II) analogue is more labile, and one of the two Zn(II) ions dissociates when free Zn(II) is less than micromolar. This raises the interesting possibility that DNA-binding affinity and therefore transcriptional regulation might depend under some circumstances on the cellular Zn(II) concentration. Whether the structural changes as a function of Zn(II) content necessary for such a postulate are a feature of GAL4 requires further investigation.

This work was supported by National Institutes of Health Grants DK09070 and GM21919. The 500-MHz NMR was supported by National Institutes of Health Grant RR03475, National Science Foundation Grant DMB-8610557, and American Cancer Society Grant RD259. This work is in partial fulfillment of the requirements for the Ph.D. degree (T.P.).

1. Oshima, Y. (1982) in *Molecular Biology of the Yeast Saccharomyces*, eds. Strathern, J., Jones, E. & Broach, J. K. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), Vol. 1, pp. 159–180.
2. Johnston, M. & Dover, J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 2401–2405.
3. Keegan, L., Gill, G. & Ptashne, M. (1986) *Science* **231**, 699–704.
4. Pan, T. & Coleman, J. E. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 3145–3149.
5. Pan, T. & Coleman, J. E. (1990) *Biochemistry* **29**, in press.
6. Pfeifer, K., Kim, K.-S., Kogan, S. & Guarante, L. (1989) *Cell* **56**, 291–301.
7. Johnston, M. (1987) *Nature (London)* **328**, 353–355.
8. Miller, J., Mclachlan, A. D. & Klug, A. (1985) *EMBO J.* **4**, 1609–1614.
9. Frey, M. H., Wagner, G., Vasak, M., Sorensen, O. W., Neuhaus, D., Wörgötter, E., Kägi, J. H., Ernst, R. R. & Wüthrich, K. (1985) *J. Am. Chem. Soc.* **107**, 6847–6851.
10. Schultze, P., Wörgötter, E., Braun, W., Wagner, G., Vasak, M., Kägi, J. H. R. & Wüthrich, K. (1988) *J. Mol. Biol.* **203**, 251–268.
11. Kägi, J. H. & Kojima, Y., eds. (1987) *Metallothionein II* (Birkhaeuser, Basel).
12. Neuhaus, D., Wagner, G., Vasak, M., Kägi, J. H. & Wüthrich, K. (1984) *Eur. J. Biochem.* **143**, 659–667.
13. Corton, J. C. & Johnston, S. A. (1989) *Nature (London)* **340**, 724–727.
14. Salmeron, J. M., Jr., & Johnston, S. A. (1986) *Nucleic Acids Res.* **14**, 7767–7781.
15. Wray, L. V., Jr., Witte, M. M., Dickson, R. C. & Riley, M. I. (1987) *Mol. Cell. Biol.* **7**, 1111–1121.
16. Kammerer, B., Guyonvarch, A. & Hubert, J. C. (1984) *J. Mol. Biol.* **180**, 239–250.